



# Inference of neuronal network spike dynamics and topology from calcium imaging data

Henry Lütcke<sup>1†</sup>, Felipe Gerhard<sup>2†</sup>, Friedemann Zenke<sup>2</sup>, Wulfram Gerstner<sup>2</sup> and Fritjof Helmchen<sup>1,3\*</sup>

<sup>1</sup> Laboratory of Neural Circuit Dynamics, Brain Research Institute, University of Zurich, Zurich, Switzerland

<sup>2</sup> School of Computer and Communication Sciences and School of Life Sciences, Brain-Mind Institute, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

<sup>3</sup> Neuroscience Center Zurich, University of Zurich and ETH Zurich, Zurich, Switzerland

## Edited by:

Peter Saggau, Baylor College of Medicine, USA

## Reviewed by:

Kresimir Josic, University of Houston, USA

Eftychios A. Pnevmatikakis, Columbia University, USA

Keith A. Cordiner, Baylor College of Medicine, USA

Keith Kelleher, Salk Institute for Biological Studies, USA

## \*Correspondence:

Fritjof Helmchen, Brain Research Institute, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland  
e-mail: [helmchen@hifo.uzh.ch](mailto:helmchen@hifo.uzh.ch)

<sup>†</sup> These authors have contributed equally to this work.

Two-photon calcium imaging enables functional analysis of neuronal circuits by inferring action potential (AP) occurrence (“spike trains”) from cellular fluorescence signals. It remains unclear how experimental parameters such as signal-to-noise ratio (SNR) and acquisition rate affect spike inference and whether additional information about network structure can be extracted. Here we present a simulation framework for quantitatively assessing how well spike dynamics and network topology can be inferred from noisy calcium imaging data. For simulated AP-evoked calcium transients in neocortical pyramidal cells, we analyzed the quality of spike inference as a function of SNR and data acquisition rate using a recently introduced peeling algorithm. Given experimentally attainable values of SNR and acquisition rate, neural spike trains could be reconstructed accurately and with up to millisecond precision. We then applied statistical neuronal network models to explore how remaining uncertainties in spike inference affect estimates of network connectivity and topological features of network organization. We define the experimental conditions suitable for inferring whether the network has a scale-free structure and determine how well hub neurons can be identified. Our findings provide a benchmark for future calcium imaging studies that aim to reliably infer neuronal network properties.

**Keywords:** calcium, action potential, reconstruction, connectivity, scale-free, hub neurons

## INTRODUCTION

Information processing in the nervous system is mediated by distributed spatiotemporal spiking activity patterns in networks of neurons. Experimentally, neuronal network dynamics have been difficult to investigate, especially under the relevant *in vivo* conditions for studying the neural underpinnings of sensory, motor, and cognitive phenomena. While multi-electrode arrays or silicon-based multi-electrode probes allow for simultaneous electrophysiological recording of spike trains from tens to hundreds of neurons with high temporal precision (Buzsaki, 2004), these techniques also suffer from a number of limitations. Assigning the recorded signal to multiple neurons in the proximity of the recording electrode remains challenging (“spike-sorting problem”) (Einevoll et al., 2011) and, most importantly, multi-electrodes sample neural tissue non-homogeneously, with highly active neurons in the vicinity of the recording electrodes being overrepresented (Olshausen and Field, 2005). This sampling bias can lead to spurious results in effective connectivity studies (Gerhard et al., 2011). Finally, extracellular multi-unit recordings commonly provide little information about cell type identity and spatial distribution of the recorded neurons.

Two-photon calcium imaging in the living brain has emerged as a powerful alternative technique, using either synthetic small-molecule or genetically-encoded calcium indicators (reviewed in Garaschuk et al., 2006; Grienberger and Konnerth, 2012; Knopfel, 2012; Looger and Griesbeck, 2012). Calcium signals imaged with high-affinity indicators can serve as proxy of spike dynamics

because each action potential (AP) is associated with a rather stereotypical somatic calcium influx causing a characteristic elementary calcium transient. Calcium imaging addresses several of the limitations inherent in multi-electrode recordings. Most importantly, it enables comprehensive sampling of the activity of many, if not all, neurons within a local population, currently up to about 500 neurons with cell number trading off against temporal resolution (1 Hz to 1 kHz) and signal-to-noise ratio (SNR) (Grewe and Helmchen, 2009; Lütcke and Helmchen, 2011). Moreover, calcium signals can be assigned unequivocally to individual neurons, permitting the analysis of the spatial distribution of neuronal activity patterns (Dombeck et al., 2009; Kampa et al., 2011) and long-term repeated functional probing of the exact same neuronal populations (Margolis et al., 2012; Lütcke et al., 2013). Finally, calcium imaging may be combined with genetic tools or *post hoc* labeling approaches to identify specific subtypes of neurons (Kerlin et al., 2010; Hofer et al., 2011; Langer and Helmchen, 2012), or with retrograde tracers to reveal long-range projection patterns of the imaged neurons (Chen et al., 2013a).

Because two-photon imaging conventionally is based on relatively slow frame rates (1–15 Hz), the majority of calcium imaging studies to date have focused on static neuronal properties such as sensory tuning curves (Ohki et al., 2005, 2006; Rothschild et al., 2010). In recent years, however, advanced laser scanning methods have been developed that enable high-speed population imaging (25 Hz and higher, up to 1 kHz) (Nikolenko et al., 2008;

Otsu et al., 2008; Grewe et al., 2010; Ranganathan and Koester, 2010; Bonin et al., 2011; Katona et al., 2012). In some cases spike times could be inferred with near-millisecond temporal precision (Grewe et al., 2010; Ranganathan and Koester, 2010; Fernández-Alfonso et al., 2013). In combination with dedicated analysis routines, high-speed two-photon calcium imaging should thus be capable, in principle, to report dynamic AP patterns in local neuronal populations. Besides providing unique opportunities to measure network activity *in vivo*, such experiments could even make it possible to extract structural information about network connectivity and topology, given sufficient accuracy of spike inference in the network.

A plethora of different algorithms have been developed to infer the spike train underlying a particular observed calcium indicator fluorescence trace. They can be broadly classified into deconvolution-based approaches to estimate changes in neuronal activity without attempting to reconstruct the occurrence of individual spikes (Yaksi and Friedrich, 2006; Vogelstein et al., 2009, 2010), template-matching techniques that infer spike times based on knowledge of the prototypical waveform of the single-AP evoked calcium transient (Kerr et al., 2007; Greenberg et al., 2008; Grewe et al., 2010; Onativia et al., 2013) and machine-learning algorithms, which require training on data sets with known “ground truth” (Sasaki et al., 2008). Little attention has been given, however, to a systematic study of how spike inference is influenced by different experimental parameters, such as SNR or signal acquisition rate. Experimentally, this is difficult to address because it would require a whole set of experiments with simultaneous *in vivo* calcium imaging and electrophysiological recordings from many neurons under various conditions. At present, only selective calibration experiments are feasible, testing the sensitivity of calcium indicators by simultaneous imaging and recording of single neurons (Kerr et al., 2005). Nonetheless, a thorough investigation of the effect of experimental parameters on spike inference would be an invaluable resource for experimentalists in order to plan experiments as well as interpret results, especially in view of the recent developments in imaging technology and indicator design. Only a few studies, employing mostly theoretical analysis or numerical simulations, have started to more systematically analyze the prospects and limits of spike inference from optical recordings (Sjulson and Miesenbock, 2007; Wilt et al., 2013) as well as of extracting network information from inferred population spike dynamics (Vogelstein et al., 2010; Mishchenko et al., 2011; Stetter et al., 2012).

Here, we present a quantitative simulation framework to generate two-photon calcium imaging signals from the spiking activity of neocortical neurons, simulated either for individual cells or for subsets of neurons within a large-scale network. Using simulated single-neuron fluorescence signals we first characterize the influence of relevant parameters of the reconstruction algorithm on spike inference, exemplified here for the recently introduced “peeling” algorithm that iteratively removes detected single-AP evoked calcium transients from the observed fluorescence signal (Grewe et al., 2010). To guide experimentalists, we provide a systematic and quantitative analysis of the impact of several parameters related either to imaging data acquisition, indicator properties, or inference routine. We evaluate

how these parameters—within value ranges relevant for real experiments—influence the fractions of correctly inferred and falsely discovered APs as well as the temporal precision of spike inference. We also extend the peeling algorithm to consider calcium indicator saturation at high spike rates.

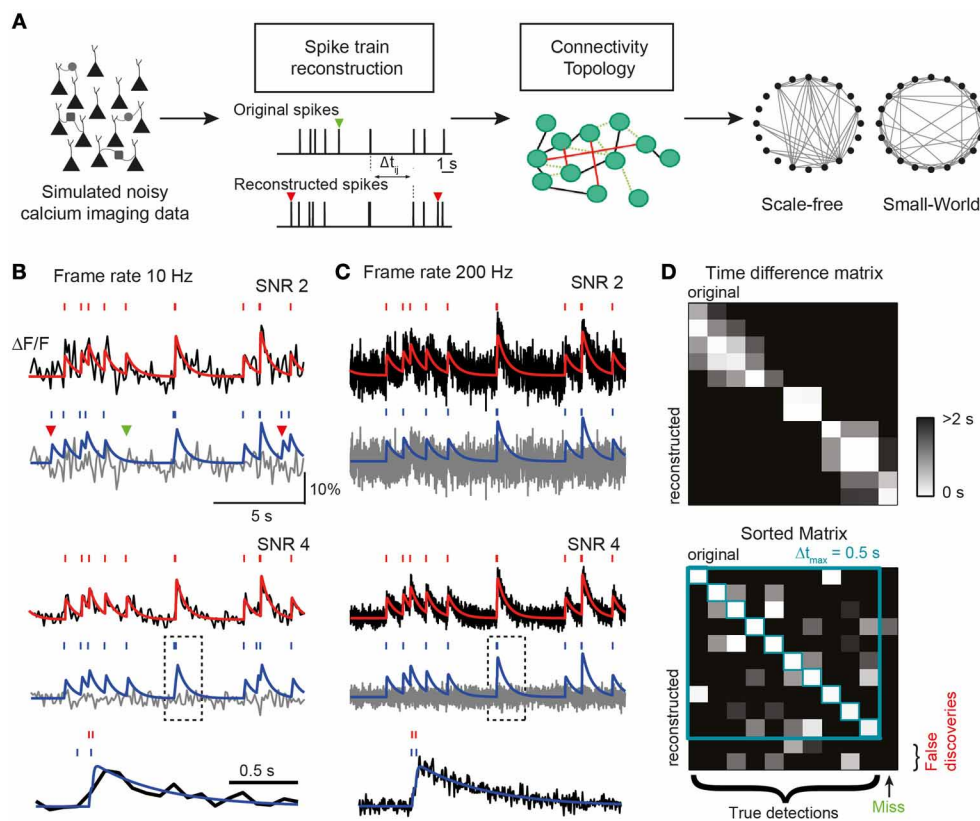
Using a large-scale neuronal network simulation, we subsequently show that structural connectivity between neurons can be partially inferred even from limited amounts of imaging data, from a sparse subset of the population, and under fluctuating, unobserved, common input. We show that parametric statistical models can extract substantially more information than pairwise cross-correlational analysis. Based on our spike inference analysis we then examine to what extent the inference of network properties is expected to deteriorate for realistic calcium imaging conditions. Finally, we investigate whether statistical network properties, such as scale-free topologies (Barabasi and Albert, 1999) and hub neurons (Feldt et al., 2011) can be recovered from the estimated connectivity matrices. Our results suggest that the current state-of-the-art in calcium imaging technology not only comes very close to the criteria required for reliable and accurate spike inference in neuronal networks but also enables at least in part to gain additional information about network connectivity and topological features.

## RESULTS

### A FRAMEWORK FOR SIMULATIONS OF NEURAL NETWORK CALCIUM IMAGING DATA

Our first aim was to create a simulation environment for mimicking actually recorded calcium indicator fluorescence traces (both in single neurons and in a network of spiking neurons), which then can be treated in exactly the same way as real experimental data. The advantages of exploring simulated fluorescence transients are: (1) the reconstructed spike trains can always be compared to the ground truth of original spike trains; (2) many artificial spike trains can be easily generated; and (3) spike trains in networks with known connectivity can be utilized to explore the possibility of extracting information about network structure from calcium imaging data. As a result, different parameters related either to the network itself, the experimental conditions or the analysis routines can be systematically varied to evaluate their relative influences (Figure 1A).

To generate artificial AP-evoked fluorescence signals, we simulated spike trains either in single neurons (Poisson process with time-independent mean firing rate  $r$ ) or in a simulated network of leaky integrate-and-fire neurons (see Materials and Methods). Spike trains were converted to fluorescence signals, taking into account their relationship to changes in intracellular free calcium concentration ( $[Ca^{2+}]_i$ ). As commonly done for experimental data, we expressed calcium signals as relative percentage fluorescence changes ( $\Delta F/F$ ). We first presumed a linear relationship between  $[Ca^{2+}]_i$  changes and  $\Delta F/F$ , which is justified for relatively isolated brief transients as they occur for sparse spiking (for treatment of non-linear indicator saturation at high firing rates see below). This is self-explanatory. In the linear case,  $\Delta F/F$  traces were generated by convolving spike trains with a kernel with a fast exponential rise (time constant  $\tau_{on}$ ) and a slower exponential decay (time constant  $\tau_{off}$ ), mimicking the stereotyped



**FIGURE 1 | (A)** Conceptual link between neuronal network dynamics and structure in our study. Network dynamics measurable with calcium imaging techniques is simulated to investigate how well spike trains can be reconstructed for a broad value range of the most important experimental parameters. Reconstruction performance is condensed in three key parameters (TPR, true positive rate; FDR, false discovery rate; and  $\sigma_{\Delta t}$ , temporal precision), which are used to analyze how well structural network properties such as topological characteristics (right) can be revealed depending on the attainable accuracy of spike train reconstruction. Missed (green) and falsely detected (red) spikes indicated by arrowheads. **(B–D)** Simulation of calcium traces from spike trains and subsequent reconstruction of firing pattern. **(B)** Top: simulated noise-free (red) and noisy (black) calcium traces for the example Poisson spike train ( $SNR = 2$ ;  $f = 10$  Hz). Bottom: reconstruction of spike train from simulated noisy calcium trace using the peeling algorithm. Reconstructed

spike train and model calcium trace in blue. Gray: residual calcium trace. Note the missed (green) and two false spikes (red). Middle: same traces but with a better SNR of 4. Bottom: expanded view of example calcium transient in gray box on faster time scale. Note the timing imprecision of reconstructed spike timing due to the low frame rate. **(C)** Same data as in **(B)** but with a higher frame rate. Note the improved reconstruction, especially at the faster time scale (bottom). **(D)** Illustration of simulated original spike train and reconstructed spike train ( $SNR = 2$ ;  $f = 10$  Hz). Spikes are matched based on the  $\Delta t$  matrix of all spike-pair-intervals  $\Delta t_{ij}$  between original and reconstructed trains. After sorting the matrix according to the rank of  $\Delta t$  and applying a threshold  $\Delta t_{max}$ , spikes are either matched or remain as “misses” or “false discoveries.” Unmatched spikes  $i$  with  $\Delta t_{ij} > \Delta t_{max}$  for all  $j$  are undetected original spikes (misses) and unmatched spikes  $j$  with  $\Delta t_{ij} > \Delta t_{max}$  for all  $i$  are spurious reconstructed spikes (false discoveries).

single AP-evoked calcium transient typically observed in neocortical pyramidal neurons with the synthetic indicator Oregon Green BAPTA-1 (OGB-1) (Kerr et al., 2005; Grewe et al., 2010). Realistic Gaussian noise was added to the simulated calcium signals to yield different SNRs (see Materials and Methods). SNR was defined as the ratio of the peak amplitude of the elementary calcium transient divided by the standard deviation of baseline activity. Finally, noisy calcium traces were subsampled from the original temporal resolution (2 kHz) to a given target frame rate,  $f$ , by selecting the center data point of each time interval ( $1/f$ ). This procedure resembles the laser scanning approach used in most two-photon microscopes. **Figures 1B,C** show examples of simulated  $\Delta F/F$  traces for two different frame rates (10 and 200 Hz) and noise levels (SNR 2 and 4).

In the following, we address three major questions with this simulation framework. First, how good is the reconstruction of spike trains in individual neurons under systematically varied conditions? Second, in how far can one extract information about physiological connectivity from the more or less accurate inferred spike times in the network? Finally, what level of reconstruction performance is necessary to infer statistical features of the underlying network topology, such as the identification of hub neurons or scale-free properties?

#### ANALYSIS OF SPIKE INFERENCE FROM SIMULATED CALCIUM SIGNALS IN INDIVIDUAL NEURONS

Our simulation framework provides a convenient strategy for the comprehensive evaluation of various algorithms that have been

devised for inferring spike trains from noisy calcium recordings (Greenberg et al., 2008; Sasaki et al., 2008; Grewe et al., 2010; Vogelstein et al., 2010). However, a comparison of different algorithms was not the goal of this study. Rather, we exemplify how the performance of one particular spike reconstruction approach, the peeling algorithm introduced in (Grewe et al., 2010), depends on different experimental parameters. In principle the same systematic approach can be followed for other reconstruction algorithms using the MATLAB code provided (Supplementary Materials).

The peeling algorithm is based on iterative subtraction of a template elementary calcium transient at event onset times detected by a Schmitt trigger routine, thus “peeling” away calcium transients until a residual noise trace remains (Figures 1B–D). Besides the parameters describing the template calcium transient, the main parameters of the peeling algorithm are two thresholds (an initial high-passing and a second low-passing one) and the minimum duration between the two threshold crossings that has to pass by in order to count as an event (see Materials and Methods). The peeling algorithm returns a list of spike times in continuous time (independent of acquisition rate). Examples of spike trains reconstructed from simulated calcium traces using the peeling algorithm are shown in Figures 1B,C. An advantage of the peeling algorithm is that it can be extended to conditions, under which indicator saturation becomes relevant (see below).

The reconstructed spike train may contain false negatives (missed spikes) and false positives (falsely discovered spikes). To quantify the performance of the inference algorithm, we compared the original and reconstructed spike train as follows. We first calculated a matrix of spike time differences ( $\Delta t$ ) for all pairs of original and reconstructed spikes (Figure 1D). We then assigned the first spike pair based on the smallest time difference, repeated this best-matching approach for the remaining spikes, and iterated until no further pair was found to meet a tolerance time window criterion (by default  $\Delta t_{\max} = 0.5$  s). The remaining “lonely” spikes constituted missed and falsely discovered spikes, respectively (Figure 1D). The outcome of this comparative approach was condensed in two main parameters: the true positive rate  $TPR_{AP}$  (number of correctly detected spikes divided by the original number of spikes; also called “sensitivity” or “recall”) and the false discovery rate  $FDR_{AP}$  (number of falsely discovered spikes divided by the number of reconstructed spikes, with  $(1 - FDR_{AP})$  also referred to as “precision”). We preferred to use the FDR rather than the false positive rate (FPR, number of time bins with falsely reconstructed spikes divided by the total number of time bins without original spike) because FPR depends on time binning and becomes arbitrarily small for high acquisition rate and sparse spiking.  $TPR_{AP}$  and  $FDR_{AP}$  on the other hand provide an intuitive quantification of the fractions of accurately and inaccurately detected spikes (see also Materials and Methods). We furthermore quantified the temporal precision of correctly retrieved spikes as the mean and standard deviation ( $\text{mean} \Delta t \pm \sigma_{\Delta t}$ ) of the difference between matched reconstructed and original spike times.

We first investigated how SNR and acquisition rate influence spike inference under conditions commonly observed for cortical

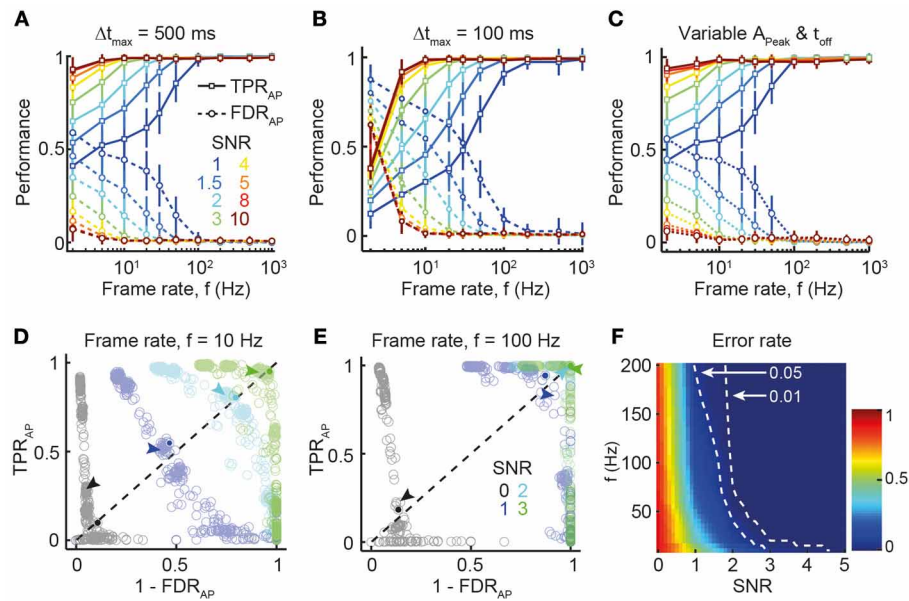
pyramidal neurons (assuming OGB-1 labeling: single-AP peak  $\Delta F/F$  amplitude 7%; decay time constant 1 s; low average firing rate of 0.2 Hz). Figure 2A summarizes the reconstruction accuracy in terms of  $TPR_{AP}$  and  $FDR_{AP}$  for different SNR levels and frame rates. As expected, lower noise levels and faster frame rates were associated with better reconstruction performance. Near-perfect reconstruction accuracy was achieved at surprisingly low frame rates, with little improvement above 30 Hz (Figure 2A), in agreement with the experimentally verified performance of the peeling algorithm  $TPR_{AP} = 95.5\%$  and  $FDR_{AP} = 1.5\%$  for 181–490 Hz acquisition rate and SNR levels between 2 and 5 (Grewe et al., 2010). Even at very high noise levels (SNR 1–1.5) of simulated traces, very good reconstruction performance was attained at higher sampling rates ( $\geq 100$  Hz). Note that choosing a shorter temporal window for declaring correct AP detections ( $\Delta t_{\max} = 0.1$  s instead of 0.5 s) impaired reconstruction accuracy at lower frame rates ( $\leq 10$  Hz) whereas accuracy at high frame rates remained largely unaffected (Figure 2B). We next asked how spike inference might be influenced by additional noise factors, such as variability in the main parameters describing the single AP-evoked calcium transient ( $A_{\text{Peak}}$  and  $\tau_{\text{off}}$ ). To address this question, we selected variable values of  $A_{\text{Peak}}$  and  $\tau_{\text{off}}$  for the calcium transient model of each AP, based on a normal distribution with mean  $A_{\text{Peak}} = 7\%$  and mean  $\tau_{\text{off}} = 1$  s and  $SD = 10\%$  of the mean value (i.e., 0.7% for  $A_{\text{Peak}}$  and 0.1 s for  $\tau_{\text{off}}$ ). Results of this simulation (Figure 2C) revealed only minor decreases in reconstruction performance, compared to the noiseless parameter scenario (Figure 2A), demonstrating that spike inference with the peeling algorithm is relatively robust against small, random fluctuations in critical parameters.

Spike reconstruction performance is not only determined by imaging parameters, such as SNR or acquisition rate, but also by properties of the detection algorithm itself, notably the settings for different thresholds (see Materials and Methods). To evaluate the effects of different decision criteria (thresholds) on signal detection, we investigated the trade-off between falsely discovered spikes ( $1 - FDR_{AP}$ ; “precision”) and correctly detected spikes ( $TPR_{AP}$ ; “recall”) [so-called “precision-recall” (PR) curve; Figures 2D,E]. Using error rate  $\alpha_{AP}$  (see Materials and Methods) as performance metric, we confirmed that sensitive spike detection is achievable at relatively low SNR, provided that frame rates are high enough (e.g.,  $> 100$  Hz; Figure 2F).

## INFLUENCE OF INDICATOR PROPERTIES AND PARAMETER CHOICES

We next investigated how properties of the calcium indicator itself might affect spike reconstruction performance. Calcium indicators are  $\text{Ca}^{2+}$ -binding molecules with characteristic binding kinetics and affinity, and they can be applied in different concentrations. These parameters affect the shape of recorded fluorescence transients, in particular the onset time, the peak fluorescence (and thus SNR), and the decay time course (Göbel and Helmchen, 2007). For the majority of commonly used calcium indicators these properties have been measured and they are usually reported for new indicators (see for example Hendel et al., 2008). We asked how





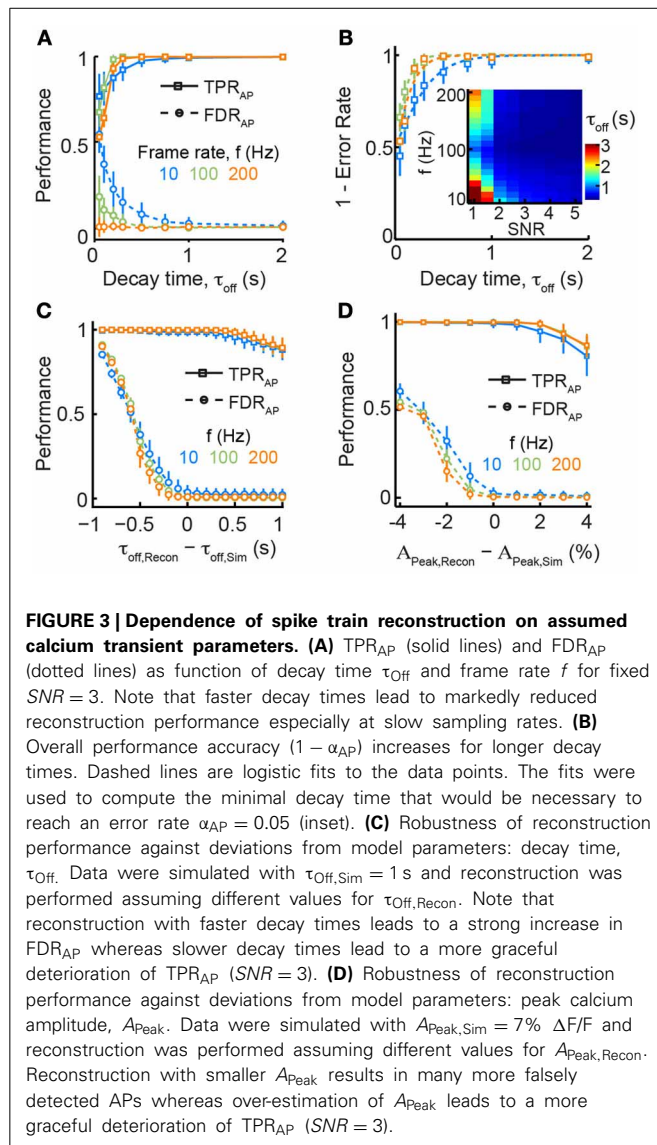
**FIGURE 2 | Dependence of spike reconstruction performance on SNR and frame rate. (A)** TPR<sub>AP</sub> (solid lines) and FDR<sub>AP</sub> (dotted lines) as function of frame rate (x-axis) and SNR (different colors). Temporal window for declaring correct AP detection: 500 ms. Mean  $\pm$  SD. **(B)** Same analysis as **(A)** but with narrower temporal window for declaring correct AP detection (100 ms). **(C)** Similar simulation as in **(A,B)**, but with variable calcium transient parameters  $A_{\text{Peak}}$  and  $\tau_{\text{off}}$ . For each AP, corresponding values for  $A_{\text{Peak}}$  and  $\tau_{\text{off}}$  were selected from a normal distribution with mean  $A_{\text{Peak}} = 7\%$  and mean  $\tau_{\text{off}} = 1$  s and  $SD = 10\%$  of the mean value. Temporal window for declaring correct AP detection: 500 ms. **(D)** PR-curve showing the trade-off between TPR<sub>AP</sub> and FDR<sub>AP</sub> for different SNR at  $f = 10$  Hz (see SNR legend in E). Data

points indicate different combinations of Schmitt-trigger thresholds for the spike reconstruction algorithm (see Materials and Methods). Arrows mark the data points corresponding to thresholds employed for reconstruction in this study. Solid circles indicate the break-even point used for quantifying overall reconstruction performance. The error rate is the normalized distance of the break-even point from the top-right corner (perfect reconstruction accuracy). **(E)** Same analysis as **(D)** at  $f = 100$  Hz frame rate. Marked data points for SNR 2 and 3 are superimposed in top-right corner, indicating near-perfect reconstruction accuracy. **(F)** Error rate for all combinations of SNR and frame rate (linear interpolation between simulated parameter combinations). Dashed lines indicated thresholds for 0.01 and 0.05 error rates.

variation in one of the most important parameters governed by indicator properties, the decay time constant  $\tau_{\text{off}}$ , affects reconstruction performance given different experimental constraints (notably frame rate and SNR). For common synthetic indicators such as OGB-1,  $\tau_{\text{off}}$  is about 0.5–1 s for typical indicator concentrations, while other recently developed highly sensitive genetically encoded calcium indicators (GECIs) can display slower decays (e.g.,  $\tau_{\text{off}} = 2$ –4 s for YC-Nano15) (Horikawa et al., 2010). Our simulations clearly show that longer decay times do not preclude accurate detection of APs, at least for sparse firing regimes (Figure 3A). Contrarily, we observed a rapid deterioration of reconstruction performance for faster decay times ( $\tau_{\text{off}} < 0.5$  s), most notably at lower acquisition rates and SNRs (Figure 3B). Intuitively, this can be explained by a too low sampling density resulting in frequent misses of the brief initial peak of calcium transients. These results suggest that the use of new calcium indicators with faster fluorescence decay times (Chen et al., 2013b) would only be beneficial in combination with imaging at sufficiently high speed.

Algorithms for spike inference from fluorescence signals typically rely on the use of a standard single-AP calcium transient (see above). Although the shape of this elementary calcium transient is rather stereotyped (as long as saturation can be neglected; see below) and often well-characterized for a certain

cell type, variations among individual cells do exist and uncertainties remain about the values to choose for the parameters of the reconstruction algorithm. We therefore explored to what extent reconstruction performance depends on the accurate choice of parameters for the elementary calcium transient. We simulated AP-evoked  $\Delta F/F$  traces with a fixed decay time constant  $\tau_{\text{off, Sim}}$  of 1 s but for spike train reconstruction we systematically varied the presumed decay time constant  $\tau_{\text{off, Recon}}$  between 0.1 and 2 s (Figure 3C). Reconstruction with shorter  $\tau_{\text{off, Recon}}$  dramatically increased the fraction of falsely detected APs while TPR<sub>AP</sub> remained unaltered. Choosing too long decay times for reconstruction led to a smaller decline of TPR<sub>AP</sub> while FDR<sub>AP</sub> remained low. Overall spike reconstruction accuracy decreased by 35% for a 50% decrease in assumed  $\tau_{\text{off, Recon}}$  ( $\tau_{\text{off, Recon}} = 0.5$  s) while a corresponding doubling ( $\tau_{\text{off, Recon}} = 2$  s) reduced accuracy by only 10% (Figure 3C). Another parameter that may be partly unknown under experimental conditions is the peak fluorescence  $A_{\text{Peak}}$  for a single AP. We found that under-estimation of  $A_{\text{Peak}}$  led to a dramatic increase in FDR<sub>AP</sub> while over-estimation again caused a more graceful degradation in TPR<sub>AP</sub> (Figure 3D). Given that the true values for  $A_{\text{Peak}}$  and  $\tau_{\text{off}}$  are frequently unknown under experimental conditions, our analysis suggests that over-estimating these parameter may be a good strategy to optimize spike reconstruction accuracy (Figures 3C,D).



### SPIKE INFERENCE UNDER CONDITIONS OF INDICATOR SATURATION

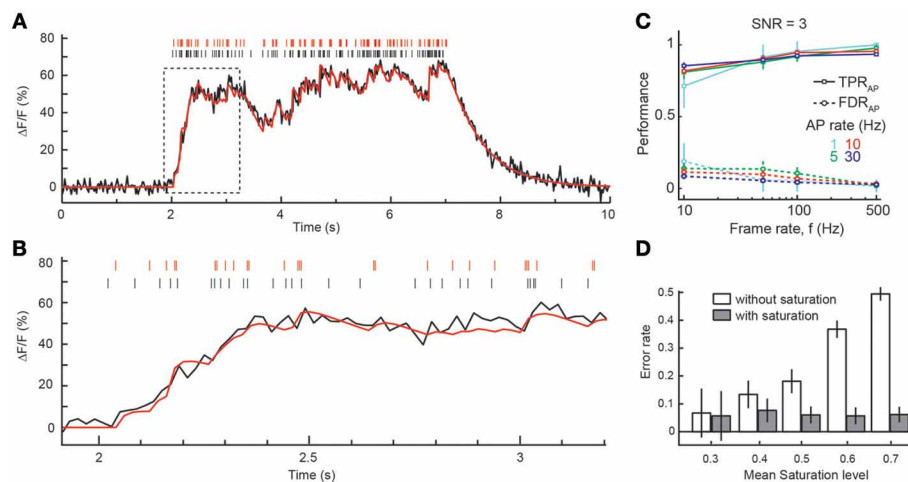
So far, we have assumed sparse spiking conditions (low firing rate), for which  $\Delta F/F$  can be presumed to relate linearly to  $[Ca^{2+}]_i$ . However, episodes of AP bursts with higher firing rates, e.g., during optimal sensory stimulation (Decharms et al., 1998) or under awake conditions (Greenberg et al., 2008), will cause larger  $[Ca^{2+}]_i$  elevations and increasingly drive high-affinity indicators into saturation. We therefore also incorporated saturation effects in our simulation framework and extended the peeling algorithm to account for saturating fluorescence transients during burst episodes of APs (see Materials and Methods). To evaluate the performance of the reconstruction algorithm for higher firing rates, we simulated saturating fluorescence traces in response to 5–10 s long episodes of AP firing rate at 1–30 Hz. At high firing rates fluorescence traces reached  $\Delta F/F$  values greater than 60% (Figure 4A), corresponding to saturation levels around 0.7. Due to saturation, the amplitude of individual AP-evoked  $\Delta F/F$  transients decreases at elevated  $[Ca^{2+}]_i$

levels (Figure 4B). However, taking this saturation effect into account in an improved version of the peeling algorithm (see Materials and Methods) we could still recover the majority of spikes, even at high firing rates, as exemplified for  $SNR = 3$  in Figure 4C. When the mean saturation level during the spiking episode increased at higher firing rates, we observed very little, if any, decrease in the overall reconstruction performance. Note, however that this result depended critically on saturation being implemented in the reconstruction step. When we simulated saturating  $[Ca^{2+}]_i$  traces but attempted to reconstruct them without taking saturation into account (using the linear approximation), error rates increased dramatically (Figure 4D). In summary, the results of the previous sections demonstrate that the peeling algorithm enables robust and highly accurate spike inference over a large range of imaging, indicator and experimental parameters.

### TEMPORAL PRECISION OF SPIKE INFERENCE

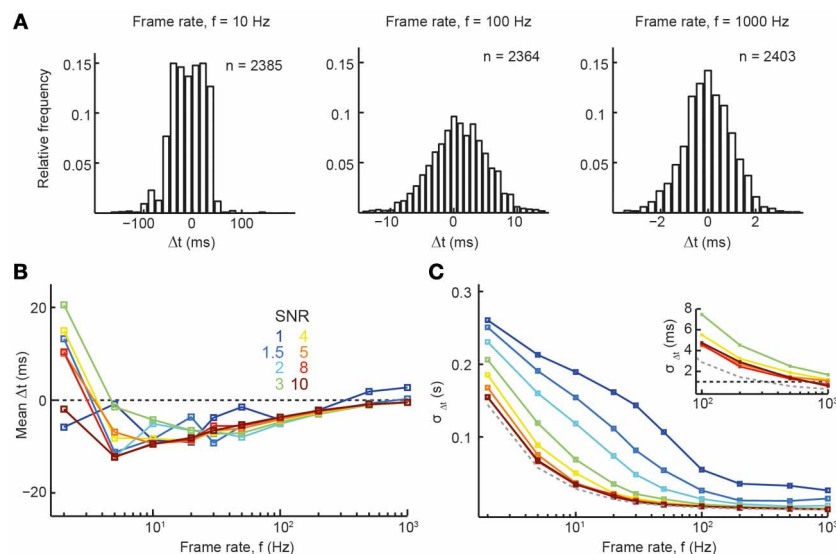
Given that accurate spike reconstruction can be achieved with our reconstruction algorithm for a wide range of parameter combinations, we next wanted to explore the temporal precision of accurate reconstructions (returning to low average firing rate of 0.2 Hz). Figure 5A shows example distributions of spike time differences at three different sampling rates (mean  $\Delta t \pm \sigma_{\Delta t}$  for 10 Hz:  $-9 \pm 35$  ms; 100 Hz:  $-4 \pm 5$  ms; 1000 Hz:  $0 \pm 1$  ms). Again, these results are in line with the experimentally determined spike time reconstruction precision of the peeling algorithm ( $SNR$  2–5; Grewe et al., 2010). At very high sampling rates and good  $SNR$ , spike train reconstruction approached sub-millisecond accuracy, as quantified by the standard deviation (SD) of the  $\Delta t$  distribution (Figures 5B,C;  $\tau_{off} = 1$  s,  $\tau_{on} = 10$  ms). For example, at a rate of 1 kHz,  $\sigma_{\Delta t}$  was 0.67 and 0.56 ms for  $SNR$  8 and 10, respectively (Figure 5C, inset), which is close to the limit set by the sampling interval. Of note, temporal precision was little influenced by  $\tau_{off}$  (data not shown) and only slightly reduced by slower onset times (at  $\tau_{on} = 20$  ms,  $\sigma_{\Delta t}$  was 0.95 and 0.76 ms for  $SNR$  8 and 10, respectively; at  $\tau_{on} = 50$  ms,  $\sigma_{\Delta t}$  was 2.90 and 3.12 ms for  $SNR$  8 and 10, respectively). These results demonstrate that recently developed high-speed imaging approaches should be adequate for accurate spike reconstruction with high temporal precision. Our analysis furthermore suggests that the recent development of high- $SNR$  GECIs (Horikawa et al., 2010; Chen et al., 2013b), even if these exhibit slightly slower onsets, may soon allow for accurate spike reconstruction from calcium imaging data with millisecond precision.

The two variables  $\alpha_{AP}$  and  $\sigma_{\Delta t}$  (ignoring mean  $\Delta t$  which will cancel out in a network reconstruction analysis based on relative differences of spike times of multiple neurons) describe in condensed form the overall performance of spike reconstruction in terms of accuracy and temporal precision. In the second part of our study, our goal was to apply this framework to simulated large-scale networks of spiking neurons with known connectivity and investigate how the attainable levels of spike reconstruction accuracy and temporal precision affect the extraction of synaptic coupling structure between neurons in order to estimate structural network connections within local neuronal populations.



**FIGURE 4 | Spike inference from episodes with high firing rates. (A)** Example simulation of an episode of 30 Hz firing for 5 s. The fluorescence trace was simulated with a frame rate of 50 Hz and  $SNR = 3$  using the model of calcium dynamics including indicator saturation. Black: simulated APs and  $\Delta F/F$  trace. Red: reconstructed APs and result of peeling algorithm with saturation model. **(B)** Zoom of initial part of the episode (boxed region in **A**). Note the decreasing amplitude of single-AP transients at high  $\Delta F/F$  due to saturation. **(C)**  $TPR_{AP}$  (solid lines) and  $FDR_{AP}$  (dotted lines) as function of frame rate (x-axis) and firing rate (different colors).  $SNR = 3$ . Temporal

window for declaring correct AP detection: 100 ms. **(D)** Dependence of error rate on mean saturation level during burst episodes with (gray) and without (white) taking saturation into account in the spike reconstruction with the peeling algorithm. Note the large increase in error rate at high saturation levels when a non-saturating, linear superposition of  $\Delta F/F$  transients is wrongly presumed. Analysis in **(D)** is based on simulations with frame rate  $\geq 50$  Hz and  $SNR = 3$  and 10. All simulated  $\Delta F/F$  traces were generated using the model of calcium dynamics including indicator saturation. All panels show mean  $\pm$  SD.



**FIGURE 5 | Precision of spike time inference. (A)** Histogram of the spike time differences between original and reconstructed spike trains for 3 different frame rates ( $SNR = 5$ ). **(B, C)** Summary parameters for the distribution of spike time differences between original and reconstructed

spikes. **(B)** Mean spike time difference,  $\Delta t$ . **(C)** Standard deviation of the distribution of spike time differences,  $\sigma_{\Delta t}$ . Dashed gray line: theoretical limit set by frame rate. Dashed black line (inset) indicates 1 ms precision. Relevant simulation parameters: firing rate = 0.2 Hz,  $A = 7\%$ ,  $\tau_{on} = 10$  ms,  $\tau_{off} = 1$  s.

## SIMULATION OF LARGE-SCALE NETWORK DYNAMICS OF SPIKING NEURONS

To investigate how the different experimental parameters impact the analysis of the network dynamics, we extended our simulation framework to generate realistic dynamics of cortical networks

under *in vivo* recording conditions. We performed large-scale neural network simulations of 25,000 neurons with sparse, random connectivity of 10% and balanced excitatory and inhibitory subpopulations (**Figures 6A,B**) in line with classical models of cortical networks (Brunel, 2000; Vogels et al., 2005). Neuronal

dynamics were modeled with leaky-integrate-and-fire (LIF) models and conductance-based synapses with individual dynamics for GABA, AMPA, and NMDA conductances (see Materials and Methods). Despite their phenomenological nature, LIF models have shown good to excellent correspondence with the complex response properties of single neurons (Badel et al., 2008; Mensi et al., 2012). In addition to the synaptic input from within the simulated population, each neuron received additional correlated Poisson spike trains whose rates were temporally modulated on all possible time scales to mimic shared input from other cortical layers and areas.

The resulting network state was balanced with irregular, asynchronous activity with sparse average firing rates of around 0.2 Hz and global rate fluctuations (see raster plot in **Figure 6A**). The explicit modeling of slow NMDA conductances (time scale  $\tau^{\text{NMDA}} = 100$  ms) led to a more decorrelated activity than in standard neural network simulations (not shown). Because of the size of the simulated network, individual synaptic couplings were relatively weak and single presynaptic APs elicited small postsynaptic potentials and did typically not trigger postsynaptic APs (**Figure 6B**).

How much can we learn about the structure and topology of the network by observing its activity and dynamics through functional signals as provided by calcium imaging experiments? Incomplete and finite observations as well as intrinsic noise sources fundamentally limit the ability to infer structural connectivity from functional signals. Experimentally, only a small fraction of all interacting circuit elements can be recorded for a limited amount of time. We respected this limitation in our simulation framework by randomly picking 50 (excitatory) neurons from the whole population and extracting their coupling structure despite the fact that there were 24,950 unobserved neurons in addition to correlated, fluctuating, external noise sources (**Figures 6A,B**). Furthermore, we did not allow arbitrarily long recording sessions, but limited ourselves to what could be observed with less than 3 h of simulated, sparse activity. These numbers were chosen to be in the realm of currently achievable experimental conditions for current high-speed calcium imaging set-ups.

#### INFERENCE OF STRUCTURAL CONNECTIVITY FROM NETWORK ACTIVITY: ESTABLISHING AN UPPER LIMIT

With the aforementioned limitations, we do not expect to unambiguously recover the network connectivity from observing network activity *even if* we had access to the unperturbed spiking activity with full temporal resolution. We therefore proceeded by first establishing an estimate of the upper bound of how well the network structure could be reconstructed in the case of infinitely precise spike time measurements. Subsequently, we analyzed how the SNR and sampling rate of calcium imaging experiments influence the reconstruction performance relative to this reference value.

To extract the coupling structure from the spike sequence of the subset of 50 neurons we used non-linear point process models (Generalized Linear Models, GLM; see Materials and Methods), which recently have been applied on electrophysiological data (Pillow et al., 2008; Vidne et al., 2012). Briefly, a probabilistic

model is fit for each neuron that explains the observed spike times using the neuron's previous activity and a constant baseline rate. Causal couplings between neurons are introduced through parameterized kernels that describe how spikes of putatively presynaptic neurons modulate the spiking probability of the modeled neuron. The model is considerably simpler than the neuron models used in the simulation, e.g., the model is unaware of the true simulation parameters and timescales and does not explicitly model the shared input from the external Poisson processes, making connectivity extraction a non-trivial task. Features of the estimated coupling filters, such as their size or statistical significance, can be used to assign a single coupling strength for each possible directed connection (Gerhard et al., 2011, 2013). After applying a threshold, we obtain an estimate of the binary connectivity structure of the network. For a given threshold, the performance of the reconstruction algorithm can be summarized in the fraction of correctly identified couplings,  $\text{TPR}_{\text{links}}$ , and the fraction of erroneously inferred connections among all detected links,  $\text{FDR}_{\text{links}}$  (**Figure 6C**). Analogous to the quantification of spike train reconstruction performance, the trade-off between the two quantities is given by the choice of threshold as visualized in a PR-curve (**Figure 6D**). In the following, we summarize the threshold-independent performance of a network reconstruction using the same measure (as in the first part on spike train inference) by the achieved error rate  $\alpha_{\text{links}}$ .

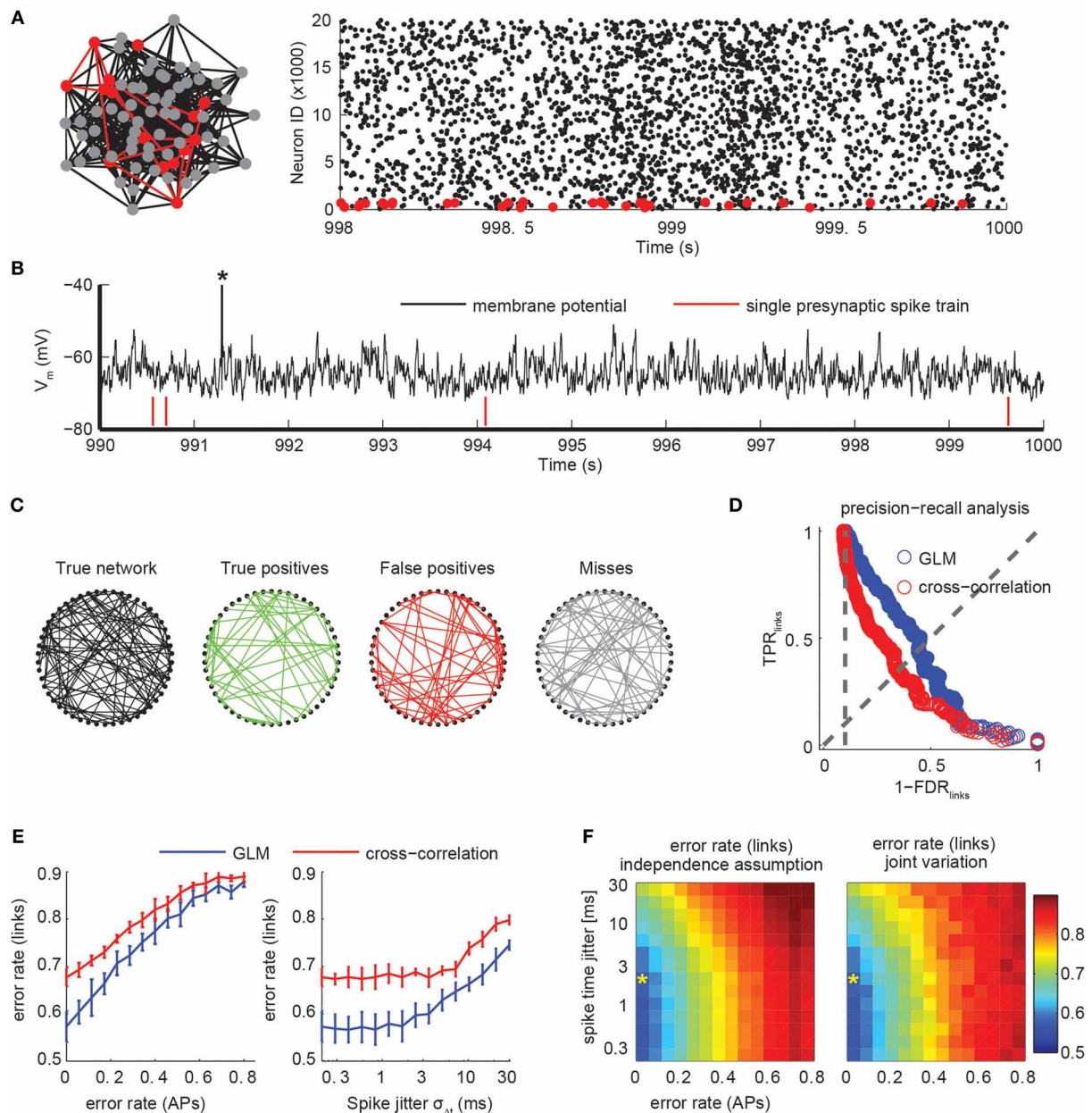
As a result, we found that with a simulated recording time of less than 3 h (or  $\sim 2000$  observed spikes per neuron), we were able to achieve an error rate between 0.5 and 0.6 (**Figure 6D**). This provides an optimal lower bound on the achievable error rate under conditions of a noiseless spike detector. We note that a naive analysis based on pairwise spike train cross-correlations resulted in significantly fewer correct links and more false positives (error rates close to 0.7, **Figure 6D**), indicating the necessity of using modern statistical models to infer network structure from functional imaging.

#### CONNECTIVITY INFERENCE AFTER IMPERFECT SPIKE TRAIN RECONSTRUCTION

An advantage of the systematic approach followed for the spike train reconstruction from noisy calcium signals is that once the key parameters  $\alpha_{\text{AP}}$  and  $\sigma_{\Delta t}$  have been determined, we can apply a surrogate transformation to the simulated spike trains of the network simulation and investigate to what degree the quality of spike train reconstruction impacts our ability to draw conclusions about the network connectivity relative to the upper bound established above. We repeated the connectivity reconstruction procedure as described above for perturbed spike trains (see Materials and Methods) and evaluated how the different performance metrics affect the connectivity inference.

First, we looked at the impact of the two parameters independently (**Figure 6E**). As expected, the measured error rate in the link reconstruction increased with increasing error rate in the AP reconstruction (**Figure 6E**, left). Given the broad range of frame rates and SNR in calcium imaging for which low error rates  $\alpha_{\text{AP}}$  were achieved (**Figure 2**), error rates  $\alpha_{\text{links}}$  close to the best value





**FIGURE 6 | Connectivity extraction using imperfect spike trains. (A)** A population of 25,000 excitatory and inhibitory neurons with sparse, random connectivity was simulated using integrate-and-fire models with conductance-based synapses. Firing activity was sparse and irregular. Subsets of 50 neurons (red) were randomly selected and only those spike trains were used to infer their mutual synaptic connectivity (left). Raster plot of the activity of the excitatory subpopulation for 2 s (right). Only every third spike is shown. **(B)** Membrane potential of a randomly selected neuron. The average neuron fired sparsely (note truncated AP marked by asterisk). Synaptic couplings were weak so that single presynaptic APs (red ticks) did not elicit APs by itself. **(C)** True network connectivity and connectivity estimated using GLMs for a randomly selected subset of 50 neurons (links divided into correctly identified links in green, false positive links in red and missed links in gray). Error rate was 0.5. **(D)** Trade-off in network reconstruction performance between  $TPR_{links}$  and  $FDR_{links}$ . Performance for unperturbed spike trains using a GLM-based reconstruction (blue) or pairwise

cross-correlation analysis (red). Error rate is defined as the intersection with the diagonal and is lower for the GLM than for the cross-correlation for all settings of the threshold. Chance level is indicated by the vertical, dotted line. **(E)** Error rates of the link reconstruction after spike perturbations for variations of the error rate in AP detection only (left, assuming no temporal jitter) and for variation of temporal precision of spike times by introducing a Gaussian jitter with width  $\sigma_{\Delta t}$  (right; assuming zero error rate in AP reconstruction). Chance level for link reconstruction is 0.9 error rate. Error bars indicate SD over repetitions using random subsets of the network. **(F)** Error rate of link reconstruction as a joint function of error rate (APs) and spike jitter, assuming both effects act independently on the error rate (left) or when jointly varied (right). The similarity indicates that effects of AP detection and its temporal precision act multiplicatively on the expected error rate for link reconstruction. Asterisk indicates the performance level that is realistically achievable with state-of-the-art high speed two-photon calcium imaging (Grewé et al., 2010; Ranganathan and Koester, 2010).

( $\sim 0.55$ ) should be reachable for many experimental conditions. With spike trains perturbed by introducing a temporal jitter  $\sigma_{\Delta t}$ , error rates in the link reconstruction increase in a more graceful manner, reaching chance level only for  $\sigma_{\Delta t} > 30$  ms (Figure 6E, right). At experimentally attainable temporal precision ( $\sigma_{\Delta t} = 1\text{--}3$  ms; see above), error rates are indistinguishable from the perfect recovery case. For comparison, we also indicated the performance reachable using a standard pairwise cross-correlational analysis, which produced substantially higher error rates than the point process models for all relevant parameter regimes (Figure 6E).

We then jointly varied  $\alpha_{AP}$  and  $\sigma_{\Delta t}$  and found that the error rate in the link reconstruction can be well-predicted by the assumption that contributions from the two factors act independently and in a multiplicative way on  $\alpha_{\text{links}}$  (Figure 6F).

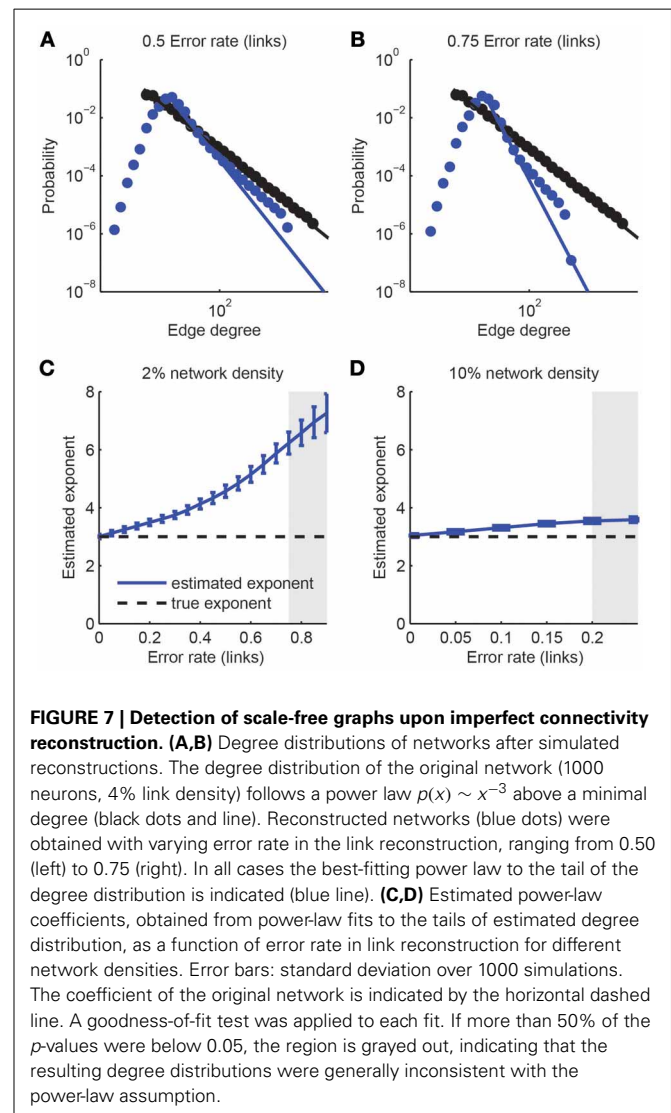
In summary, our analysis indicates that GLMs allow at least partial extraction of effective network connectivity from imperfectly reconstructed spike trains of moderate length from spontaneous, asynchronous network activity. Importantly, we show that a temporal precision of spike reconstruction in the millisecond-range is the major determinant for accurate estimation of neuronal couplings, given that near-perfect reconstruction accuracy can be achieved under a wide range of imaging conditions ( $\alpha_{AP} \sim 0$ , see above). This finding further highlights the importance of recently developed high-speed imaging approaches.

### IDENTIFICATION OF GRAPH TOPOLOGY

Through our simulation framework, we estimate that noisy calcium signals from a set of neurons immersed in a larger cortical neural population contain significant information about the underlying network structure that goes beyond what can be trivially extracted using cross-correlational analysis. Absolute reconstruction performance is, however, currently limited by several factors, some of which are at least partially under the control of the experimentalist. We therefore asked what level of reconstruction will be sufficient for inferring high-level features and statistical properties of the network structure. We addressed this question with a sensitivity analysis for two different applications that are inspired by recent experimental and theoretical studies: the quantification of scale-free properties of the network and the identification of hub neurons with imperfect network reconstruction.

First, we considered the class of *scale-free networks* (Barabasi and Albert, 1999): Scale-free networks are characterized through a degree distribution that follows a power law  $p(x) \sim x^{-\mu}$ , where the probability that a neuron of the network has  $x$  synaptic connections scales like a power law with characteristic exponent  $\mu$ . We generated prototypical examples of scale-free networks of size  $N = 1000$  neurons and exponent  $\mu = 3$  and analyzed the impact of expected network reconstruction performance with varying error rates  $\alpha_{\text{links}}$  and assuming that the identification of a link between any two neurons is statistically independent of any other connection (Figure 7). Thus, we obtained simulated reconstructed networks that differed in their statistical properties from the original scale-free network. We fitted power-law distributions to the reconstructed degree distributions and assessed

goodness-of-fit with a semi-parametric bootstrapping method (Figures 7A,B). Although link over- and under-sampling does in general not result in a pure power-law (Han et al., 2005; Stumpf et al., 2005), many of the resulting networks were still compatible with the scale-free assumption for the tail of their degree distributions ( $p_{\text{gof}} > 0.05$ ). However, the estimated exponent could be considerably different from the original one. In general, we found a stronger decay of the degree distribution (more negative power law exponent) upon imperfect network reconstruction (Figures 7C,D), regardless of the underlying network density (varied between 2 and 10%), i.e., the number of links present in the network. Only if reconstruction errors get moderately high (e.g.,  $\alpha_{\text{links}} > 0.75$  for networks with 2% link density), the power-law distribution was not evident anymore. The difference between the fitted and original scaling exponent varied with the network density. For denser networks, the sensitivity to imperfections of the detection of the scale-free property based on the data sample is increased. For example, an error



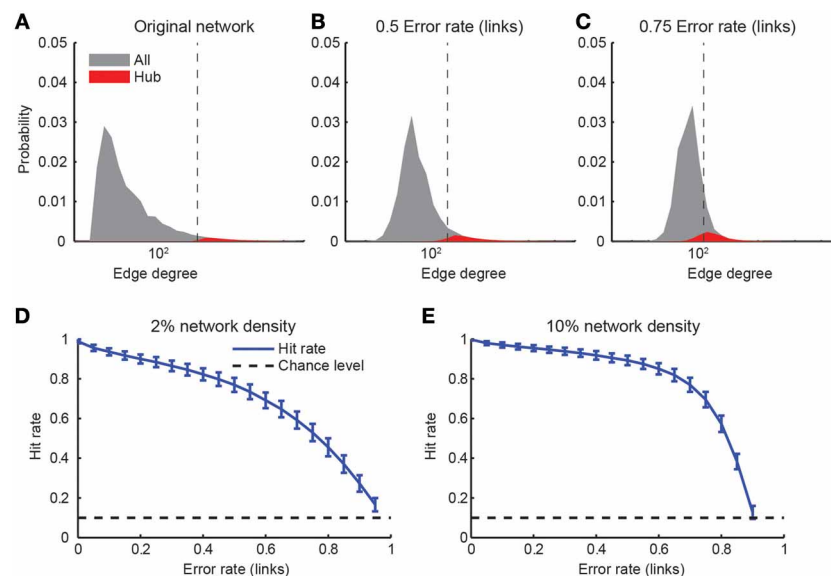
rate of up to 0.75 was tolerated for sparsely connected networks (density 2%) while an error rate of 0.2 (which is not realistically achievable under experimental conditions) was sufficient to make the degree distribution distinct from a power-law for denser networks (10% connectivity; **Figure 7D**). A detailed analysis of the contributions separate from missed links and false positive links indicated that the consistent over-estimation of the power-law exponent was mostly due to the introduction of false positive links (results not shown) while it remained almost unaffected by non-perfect  $\text{TPR}_{\text{links}}$ . This suggests that when considering the trade-off between  $\text{TPR}_{\text{links}}$  and  $\text{FDR}_{\text{links}}$ , emphasis should be put to minimize the number of false positive links on the expense of the link detection power.

**Hub neurons** are another concept relevant for graph theory (Feldt et al., 2011): Hub neurons are those neurons with a comparatively high degree, i.e., large numbers of incoming or outgoing connections in the network. As before, we started with networks of size  $N = 1000$  neurons and a scale-free topology. We classified individual neurons as hub neurons when their degree was in the upper-most decile of the degree distribution (**Figure 8A**). Imperfect network reconstruction, for example due to non-perfect link reconstruction or the insertion of false positive links, generally scattered the degree distribution and therefore the identity of hub neurons. Consequently, not all original hub neurons remained within the top decile of the estimated degree distribution (**Figures 8B,C**). We quantified the robustness with regard to reconstruction errors with the “hit rate,” i.e., the fraction of hub neurons of the original network that were correctly classified as hub neurons in the reconstructed network.

As expected, the hit rate decreased steadily with increasing error rate in the link reconstruction (**Figures 8D,E**), with only slight dependence on the original network density. Surprisingly, however, hub classification could still be robustly achieved (75% hit rate) for relatively high error rates (0.6–0.75). We conclude that hubs can be detected in a relatively robust manner even in the presence of large link reconstruction errors.

## DISCUSSION

In this study we analyzed inference of spike dynamics in neuronal networks, as well as inference of underlying structural properties, based on population fluorescence data as they are typically acquired during *in vivo* two-photon calcium imaging experiments. We established a simulation framework that—unlike experiments—allows comprehensive exploration of parameter spaces. Systematic parameter variation helps to explore the limits of what is currently achievable, identify critical parameters, motivate further methods improvements, and guide the experimenter in the optimization of their imaging conditions. Our results indicate that with state-of-the-art methods, especially high-speed two-photon calcium imaging, it is now feasible to reconstruct spike trains in populations of several tens of neurons with high precision. Remaining uncertainties in exact spike times lower, but do not preclude, retrieval of partial information about network connectivity and topological features. *In vivo* calcium imaging combined with the analysis tools described here thus promises to become a powerful method to analyze the functional organization of neuronal networks in the brain.



**FIGURE 8 | Detection of hub neurons upon imperfect connectivity reconstruction.** (A–C) Degree distributions of networks after simulated reconstructions. The degree distribution of the original network (A) followed a power law  $p(x) \sim x^{-3}$  above a minimal degree of 20. Hub neurons were defined as the 10% of neurons with the highest degree (red areas). Imperfectly reconstructed networks were obtained

by assuming varying error rates (B,C). Vertical dashed lines indicate the beginning of the upper-most decile of the estimated distribution. (D,E) Hit rate of hub neuron identification as a function of varying degree of link reconstruction error for two different network densities. Error bars: standard deviation over 1000 simulations. Chance level indicated by dashed lines.

## POTENTIAL AND LIMITS OF SPIKE INFERENCE

In the first part of this study, we highlighted the utility of our simulation framework to reveal non-trivial relations among the key imaging parameters and the accuracy of spike inference. Our systematic analysis of spike train reconstruction (here exemplified for the peeling algorithm) provides a resource for experimentalists, from which the expected reconstruction performance for a given set of experimental conditions can be obtained (**Figures 2–5**). Similar analyses can easily be implemented for alternative spike reconstruction algorithms (Yaksi and Friedrich, 2006; Vogelstein et al., 2009, 2010) and other experimental conditions. Our analysis reveals a number of novel insights. First, we show that near-optimal spike train reconstruction (i.e., detecting nearly all spikes with negligible numbers of false positives) may be achieved at experimentally tractable noise levels ( $\text{SNR} \geq 2$ ) with surprisingly low sampling rates (20–30 Hz). Frame scanning with conventional mirror-based laser scanners typically is limited to 10–20 Hz for 50–100 lines, as the standard galvanometer can be driven at about 1 kHz at maximum. Nonetheless, faster acquisition rates (even  $>100$  Hz) are possible with standard scanners using free line-scans on preselected neuronal subsets (Göbel and Helmchen, 2007; Nikolenko et al., 2007; Lillis et al., 2008; Rothschild et al., 2010), albeit usually at the expense of total dwell time per cell (and thus SNR). Alternatively, full-frame scanning up to 100 Hz has been achieved by fast scanning along one axis with a resonant galvanometer (4–12 kHz resonance frequencies) (Rocheffort et al., 2009; Bonin et al., 2011). All these methods with the capability of scanning neuronal populations at greater than video rate (25 Hz) should enable high-quality spike inference.

Two trade-offs, however, need to be considered. Gaining speed in the range of 10–100 Hz will only help to improve spike inference if a sufficient SNR is maintained (**Figure 2A**). In addition, effective sampling rate usually relates inversely to the number of recorded cells for a given SNR, so that a compromise between speed and population size is required. For a fair comparison of imaging approaches and spike reconstruction accuracies one should thus rely on populations of similar size. Fundamentally, these trade-offs arise from the fact that SNR is ultimately limited by photon statistics (Wilt et al., 2013). Of course, the SNR will be lower at higher frame rates given the same excitation power. However, high-speed scanning at low excitation rate can have the additional benefit of reduced phototoxicity and thus prolonged experiment time (Chen et al., 2012). In addition, detection of fluorescence photons should be maximized, for example by using a low-magnification, high numerical aperture objective for detection (Oheim et al., 2001) or by employing supplementary detection schemes (Engelbrecht et al., 2009).

A second insight is that ultra-fast imaging (sampling rates  $>500$  Hz) in combination with high SNR levels permits spike reconstruction with millisecond or even sub-millisecond precision. Such high acquisition rates for neuronal populations, e.g., 0.5 kHz for about 50 neurons, are possible with random-access scanning using acousto-optical deflectors (Reddy and Saggau, 2005; Grewe et al., 2010; Ranganathan and Koester, 2010). To fully exploit the potential of highest-speed calcium imaging it will be essential, however, to reduce additional noise sources such as

baseline fluctuations, bleaching effects, or motion artifacts to a minimum.

Our analysis furthermore suggests that the combination of ultra-fast imaging and high SNR may be achievable with the next generation of highly sensitive GECIs (Horikawa et al., 2010; Chen et al., 2013b). Our results reveal that faster decay kinetics of the indicator does not necessarily lead to better spike reconstruction. Intuitively, if the decay time of the calcium indicator dye is faster than the frame duration, peaks will occasionally be missed, thereby reducing detection accuracy. Thus, emerging faster calcium indicators, especially new GECIs, might be of limited use unless combined with new microscopy techniques that allow for faster image acquisition. In addition, because slow onset kinetics slightly reduces the achievable temporal precision, GECIs with fast onset characteristics are desirable if uncertainty of spike times needs to be minimized, for example to extract network structural properties.

## INFERRING NETWORK STRUCTURE FROM CALCIUM IMAGING DATA

Our simulation framework summarizes the effects of noise, calcium indicator dynamics, and imperfect reconstruction with two key quantities: the error rate (a combination of the fraction of correctly identified spikes and the rate of erroneously detected spikes) and the precision of reconstructed spike times. Based on these parameters, we could estimate how well we are likely to be able to estimate network connectivity. The goodness of connectivity estimation may again be summarized by the same key metric: the error rate in the network link reconstruction. This parameter has an influence on the estimation of graph properties, as we have exemplified for scale-free topology and the identification of hub neurons. Similarly, our framework should allow analysis of the robustness of other graph statistics such as small-world properties or the distribution of higher-order network motifs.

The modular structure of our framework enabled us to predict the effect of experimental parameters on the global estimation of statistical graph properties by simply following its effect through the different stages. For the benefit of experimenters, we can illustrate the utility of our simulation framework with a practical example, assuming state-of-the-art technology. Using the high-affinity calcium indicator OGB-1 in combination with AOD-based random-access sampling at  $\approx 500$  Hz, neuronal calcium signals have been measured in mouse neocortex with SNR up to 5 (Grewe et al., 2010). These parameters will allow spike reconstruction with near-perfect accuracy (**Figure 3A**,  $\text{TPR}_{\text{AP}} > 0.95$ ,  $\text{FDR}_{\text{AP}} < 0.05$ , therefore,  $\alpha_{\text{AP}} < 0.05$ ) as well as 2–3 ms temporal precision (**Figure 5C**, inset). In this parameter range, spike detection is sufficiently accurate so that no information loss is expected to occur with respect to the reconstruction of synaptic connections (**Figure 6E**, left). The imperfect temporal precision, however, leads to a slightly increased error rate in the link reconstruction compared to sub-millisecond temporal precision (**Figure 6E**, right). The approximate error rate can be also directly read from **Figure 6F**. Under otherwise optimal conditions, we would obtain an error rate  $\alpha_{\text{links}} \approx 0.60$ , therefore we expect to recover around 40% of the synaptic connections ( $\text{TPR}_{\text{links}} = 0.4$ ) with  $\text{FDR}_{\text{links}} = 0.6$ . According to this estimate, robust identification of hub neurons should still be feasible (**Figures 8D,E**) and



scale-free properties might be just identifiable (but heavily biased) depending on the underlying connection density (**Figures 7C,D**).

The absolute numbers in the example above should be regarded as ballpark estimates rather than precise predictions of reconstruction performance because inference is based purely on simulated data and our particular choice of algorithms. Real reconstruction performance could be weaker than predicted due to effects such as unobserved neuromodulation, weak synaptic strengths, or oscillatory background activity. In addition, connectivity reconstruction could potentially be improved by using more complex point process models that explicitly model global state fluctuations (Smith and Brown, 2003) by attempting to infer the dynamics of unobserved neurons (Vidne et al., 2012) or by employing Bayesian methods (see below).

Few other studies explored the possibility of reconstructing network connectivity from calcium imaging data. Mishchenko and colleagues presented a sensitivity analysis using a Bayesian approach combined with MCMC (Markov Chain Monte Carlo) techniques (Mishchenko et al., 2011). The combined estimation of spike times and connectivity make their approach computationally very expensive. Our modular approach offers the advantage that we can identify the crucial experimental parameters and propagate their effect through the spike time estimation to the network level. Mishchenko et al. did not consider the effect of imperfect link reconstruction for the inference of higher-level topological features of the network. Our results suggest that the optimal choice of experimental parameters can strongly depend on which feature of the network one wants to estimate most reliably. Finally, we note that their recommendation to use frame rates of at least 30 Hz “to achieve meaningful reconstruction results” (Mishchenko et al., 2011) is in alignment with our findings, although both methodology and details of the simulation vary considerably between approaches.

Another study proposed a method based on information-theoretic measures to infer effective connectivity from calcium imaging experiments and evaluate it on simulated data (Stetter et al., 2012). Our approach extends their analysis in a number of different aspects. First, their approach is only suitable for recordings with low sampling frequencies. The information-theoretic measure (transfer entropy) they use to infer couplings does not easily scale up to high-speed recordings because of the need to estimate high-dimensional probability distributions. Activity levels were coarsely discretized, which will have a negative impact on the performance especially for low SNRs. Second, their measure of coupling strength cannot distinguish between excitatory and inhibitory couplings, is pairwise only (i.e., it does not take into account the activity of other recorded cells) and is limited in how temporal aspects of couplings between cells can be modeled. All of these limitations can be overcome by the use of non-linear point process models based on the estimated spike trains, as proposed in this study. Last, (Stetter et al., 2012) did not study the impact of different experimental conditions and can therefore give only limited guidance on how to design experimental set-ups.

We note that we use the peeling algorithm coupled with point process models as an example of a combination of methods to extract connectivity. We favored a two-step procedure of first reconstructing spike trains and then inferring connectivity

based on estimated spike times over directly modeling couplings between fluorescence traces (Stetter et al., 2012; Turaga et al., *in press*). Given that spikes are triggering neurotransmitter release at synapses (at least for most cortical cell types) we expect our approach to be closer to the biological mechanism of how neural signals are coupled, and therefore to be superior in estimating connectivity. A formal comparison would though require the evaluation of both methods on the same simulated data sets.

Due to our modular analysis, the last part of our study (how high-level network properties can be recovered given an expected link reconstruction performance) is independent of the underlying methods to infer connectivity once the performance of any such reconstruction method is quantified in terms of expected link detection power and false discovery rates. Thus, a similar sensitivity analysis could be performed for additional network measures (such as small-world properties or network motifs), and our conclusions readily apply to connectivity estimates obtained from electrophysiology, e.g., from multi-electrode arrays or other imaging modalities.

## FUTURE DIRECTIONS

Two-photon calcium imaging has conventionally been calibrated by simultaneous imaging and electrophysiological recording of single neurons (Kerr et al., 2005). Based on the ground truth provided by the electrical recording, the performance of spike inference from calcium imaging data can be verified, albeit the extent of such analysis (e.g., to test various experimental conditions) is limited due to the technical difficulties. Similarly, connectivity inference and extraction of topological network properties will eventually require experimental verification against ground truth data. A first attempt has been made by Gerhard et al. (2013) who showed that the effective connectivity derived from spiking activity using a point process model similar to the one used here matches the physiological connectivity in a very small, but well-defined neural circuit. While it remains a difficult task to test these methods on larger populations, novel approaches have recently emerged that at least partially may allow such verification, including large-scale anatomical circuit reconstructions using electron microscopy (Bock et al., 2011; Briggman et al., 2011) and automated light-microscope techniques in combination with expression of cell type-specific markers and trans-synaptic tracers (Osten and Margrie, 2013). In addition, connectivity mapping can be performed following *in vivo* calcium imaging and re-identification of the recorded neuronal populations in extracted tissue using various physiological techniques, such as multi-cell electrophysiological recordings in acute brain slices (Hofer et al., 2011; Ko et al., 2011) or two-photon photo-stimulation with single-cell resolution using caged compounds or specially tailored opsins (Prakash et al., 2012). At present, these methods are, however, limited to specific individual neurons or small groups of neurons at most.

Our study may be extended in several directions. In particular, the heterogeneity of neuronal cell types could be taken into account. All above considerations straightforwardly apply to superficial neocortical pyramidal neurons, which produce large single-AP evoked calcium transients and display relatively low

spontaneous firing rates. Our extension of the peeling algorithm to account for indicator saturation should also allow reconstruction of brief AP bursts and episodes of higher firing rates, which is especially relevant for awake studies (Greenberg et al., 2008; Wolfe et al., 2010) and for deep-layer cortical pyramidal neurons that generally display higher AP rates (De Kock and Sakmann, 2008). Inhibitory interneurons, especially fast-spiking parvalbumin-expressing cells, have much smaller single-AP evoked calcium transients as well as higher firing rates (Hofer et al., 2011), suggesting that accurate spike inference may not be feasible for these neurons (SNR presumably below 0.5). On the other hand, recent *in vitro* electrophysiological work indicated that inhibitory neurons form a relatively unspecific and densely connected network in neocortical circuits (Fino and Yuste, 2011; Packer and Yuste, 2011). Another promising direction of future calcium imaging studies is to resolve the precise functional and structural topology of highly specific local networks of pyramidal neurons (Song et al., 2005). Our findings indicate that the technical requirements to achieve this goal may be just in reach.

## MATERIALS AND METHODS

### SIMULATION OF SINGLE-NEURON SPIKE TRAINS, CALCIUM DYNAMICS, AND INDICATOR FLUORESCENCE SIGNALS

Simulation of single-neuron spike trains and calcium indicator fluorescence signals and all analysis were performed in Matlab (The Mathworks, Natick, MA, USA). We generated spike trains by a Poisson process assuming a low mean firing rate (0.2 Hz) similar to what has been reported for spontaneous activity of pyramidal cells in both anesthetized and awake rodent sensory cortex (Wolfe et al., 2010). In addition, to examine the effect of calcium indicator saturation we explored episodes of higher firing rates between 1 and 30 Hz as they occur in pyramidal neurons, e.g., upon sensory stimulation (Greenberg et al., 2008).

A general description of AP-evoked fluorescence signals needs to consider the transformation of changes in intracellular free calcium concentration  $[Ca^{2+}]_i$  to the particular type of fluorescence readout. Here, we use the widely adopted  $\Delta F/F$  approach, expressing calcium signals as relative percentage fluorescence changes after background subtraction. In this case the transformation between  $[Ca^{2+}]_i$  and fluorescence signal is given by (Helmchen, 2012):

$$[Ca^{2+}]_i = \frac{[Ca^{2+}]_{rest} + K_d \frac{\Delta F/F}{\Delta F/F_{max}}}{\left(1 - \frac{\Delta F/F}{\Delta F/F_{max}}\right)} \quad (1)$$

or reversely expressed by:

$$\Delta F/F = \Delta F/F_{max} \frac{[Ca^{2+}]_i - [Ca^{2+}]_{rest}}{[Ca^{2+}]_i + K_d} \quad (2)$$

Here,  $[Ca^{2+}]_{rest}$  denotes the resting calcium concentration,  $K_d$  the dissociation constant of the calcium indicator, and  $\Delta F/F_{max}$  the maximal  $\Delta F/F$  reached upon saturation. Note that this transformation is a non-linear relationship. For fluorescence transients far from saturation ( $[Ca^{2+}]_i \ll K_d$ ) Equation 2 can be linearized to:

$$\Delta F/F = \frac{\Delta F/F_{max}}{K_d} ([Ca^{2+}]_i - [Ca^{2+}]_{rest}) = \frac{\Delta F/F_{max}}{K_d} \Delta [Ca^{2+}]_i \quad (3)$$

This linear description is a good approximation for AP-evoked fluorescence signals in the low firing regime measured for example with a high-affinity indicator such as OGB-1 (Grewe et al., 2010) (Figures 9A,B). In this case each AP evokes a stereotypic, elementary somatic calcium transient, which can be approximated with a rapidly rising and exponentially decaying function:

$$\Delta F/F = A \left(1 - e^{-(t-t_0)/\tau_{on}}\right) e^{-(t-t_0)/\tau_{off}} \text{ for } t \geq t_0 \quad (4)$$

Here,  $t_0$  denotes the time point of spike occurrence,  $\tau_{on}$  the onset rise time,  $\tau_{off}$  the decay time, and  $A$  an amplitude scale parameter. The peak amplitude  $A_{peak}$  of the single-AP evoked calcium transient is given by:

$$A_{peak} = A \tau_{off} \left(\frac{\tau_{on}}{\tau_{on} + \tau_{off}}\right)^{\frac{\tau_{on}}{\tau_{on} + \tau_{off}}} (\tau_{on} + \tau_{off})^{-1} \quad (5)$$

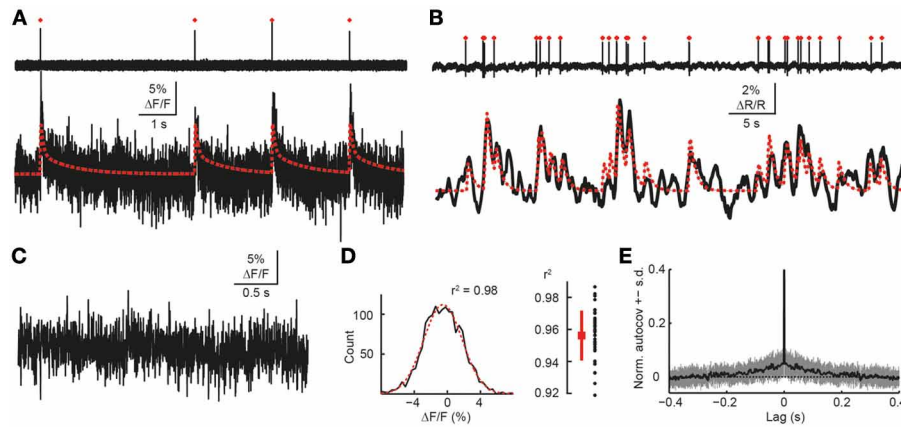
For the calcium indicator OGB-1, typical values of these parameters for neocortical pyramidal neurons are  $\tau_{on} = 10$  ms,  $A_{peak} = 7\%$   $\Delta F/F$ ,  $\tau_{off} = 0.5$ – $1$  s (Grewe et al., 2010). For the low firing regime we used the canonical elementary  $\Delta F/F$  transient (Equation 4) as impulse response function. Other more complex shapes of the elementary transient, for example a double-exponential decay (Grewe et al., 2010), could be easily incorporated into the simulation framework. Because of the linear approximation, we obtained the fluorescence traces for the entire duration of the simulations by convolving the simulated spike trains with this elementary  $\Delta F/F$  transient.

At higher AP firing rates,  $[Ca^{2+}]_i$  may reach levels sufficiently high to cause substantial saturation of the calcium indicator. We therefore incorporated the possibility to account for indicator saturation in our simulation framework. Assuming a non-cooperative calcium binding characteristics, the saturation level  $S$  (ranging from 0 to 1) is given by:

$$S = \frac{[CaB]}{[B]_T} = \frac{[Ca^{2+}]_i}{[Ca^{2+}]_i + K_d} = \frac{[Ca^{2+}]_{rest} + K_d \frac{\Delta F/F}{\Delta F/F_{max}}}{[Ca^{2+}]_{rest} + K_d} \quad (6)$$

Here,  $[B]_T$  denotes the indicator concentration in the cell and the equation's right side was obtained by insertion of equation 1. Importantly, indicator saturation not only leads to a non-linear transformation between  $[Ca^{2+}]_i$  and  $\Delta F/F$  but also directly affects buffered  $[Ca^{2+}]_i$  dynamics, an aspect that has been neglected in previous attempts to incorporate indicator saturation in spike inference algorithms (Vogelstein et al., 2009; Stetter et al., 2012). Differentiation of Equation 6 with respect to  $[Ca^{2+}]_i$  yields the so-called  $Ca^{2+}$ -binding ratio  $\kappa_B$  (or “buffering capacity”) of the indicator, which decreases with increasing  $[Ca^{2+}]_i$  levels near saturation:

$$\kappa_B = [B]_T \frac{\partial S}{\partial [Ca^{2+}]_i} = \frac{\partial [CaB]}{\partial [Ca^{2+}]_i} = \frac{[B]_T K_d}{([Ca^{2+}]_i + K_d)^2} \quad (7)$$



**FIGURE 9 | Validation of simulation framework with experimental data.**

**(A)** Simultaneous cell-attached recording and high-speed two-photon calcium imaging in mouse neocortex *in vivo*. Top trace: cell-attached recording, APs marked by red dots. Bottom trace: measured cellular  $\Delta F/F$  calcium signal (black) as well as simulated  $\Delta F/F$  traces (red) for the recorded spike train. Imaging data were acquired with OGB-1 at 490 Hz sampling rate (Grewe et al., 2010). Note that a non-saturating model with double-exponential decay was used in this case to generate the simulated trace. **(B)** Simultaneous cell-attached recording and two-photon calcium imaging using the genetically-encoded calcium indicator YC3.60 (Lütcke et al., 2010). Top trace: cell-attached recording, APs marked by red dots.

Bottom trace: measured cellular calcium signal (black) as well as simulated  $\Delta F/F$  traces (red) for the recorded spike train (sampling rate: 781 Hz; expressed as relative percentage change  $\Delta R/R$  of the YFP/CFP fluorescence ratio). A non-saturating model with single-exponential decay was used to generate the simulated trace. **(C)** Noise  $\Delta F/F$  trace during episode without AP, as confirmed by simultaneous cell-attached recording (data not shown). Sampling rate: 490 Hz. **(D)** Left: distribution of signal intensities for data shown in **(C)**. Gaussian fit in red ( $r^2 = 0.98$ ). Right: distribution of goodness-of-fit of Gaussian fits ( $r^2$ ) for pooled data set (96 s total recording time). **(E)** Mean normalized autocovariance ( $\pm$ SD) for pooled noise data set. Peak at 0 s lag clipped.

Note that the  $\text{Ca}^{2+}$ -binding ratio critically depends on the indicator's  $\text{Ca}^{2+}$ -binding affinity and its total concentration. The effect of adding an exogenous  $\text{Ca}^{2+}$ -buffer such as the indicator on AP-evoked somatic calcium signals is well-understood for neocortical pyramidal neurons and is typically approximated by a single-compartment model, which assumes chemical equilibrium and neglects diffusion (Helmchen and Tank, 2011). The model additionally considers an endogenous  $\text{Ca}^{2+}$ -binding ratio  $\kappa_S$ , which we assumed to be constant [ $\kappa_S = 100$ ; (Helmchen et al., 1996)], and the  $\text{Ca}^{2+}$  extrusion rate  $\gamma$  ( $800 \text{ s}^{-1}$ ) (Helmchen and Tank, 2011).  $[\text{Ca}^{2+}]_{\text{rest}}$  was assumed 50 nM. The relaxation of  $[\text{Ca}^{2+}]_i$  from an elevated level back to resting level is then described by the following non-linear differential equation:

$$\begin{aligned} \frac{d[\text{Ca}^{2+}]_i}{dt} &= \frac{-\gamma \Delta[\text{Ca}^{2+}]_i}{(1 + \kappa_S + \kappa_B)} \\ &= -\gamma \left( 1 + \kappa_S + \frac{[\text{B}]_T K_d}{([\text{Ca}^{2+}]_i + K_d)^2} \right)^{-1} ([\text{Ca}^{2+}]_i - [\text{Ca}^{2+}]_{\text{rest}}) \end{aligned} \quad (8)$$

To calculate the model  $[\text{Ca}^{2+}]_i$  traces for a given spike train, we numerically solved Equation 8 for each spike-to-spike interval starting from the  $[\text{Ca}^{2+}]_i$  level reached after each AP. This level was calculated by incrementing the pre-AP  $[\text{Ca}^{2+}]_i$  level at the moment of the next spike's occurrence  $t_{\text{spike}}$  by

$$\begin{aligned} \Delta[\text{Ca}^{2+}]_i(t_{\text{spike}}) &= \frac{\Delta[\text{Ca}^{2+}]_T}{(1 + \kappa_S + \kappa_B)} \\ &= \left( 1 + \kappa_S + \frac{[\text{B}]_T K_d}{([\text{Ca}^{2+}]_i(t_{\text{spike}}) + K_d)^2} \right)^{-1} \Delta[\text{Ca}^{2+}]_T \end{aligned} \quad (9)$$

Here,  $\Delta[\text{Ca}^{2+}]_T$  denotes the total intracellular calcium concentration change caused by an AP, which was assumed  $7.6 \mu\text{M}$ . The reduction of  $\kappa_B$  at elevated  $[\text{Ca}^{2+}]_i$  levels due to indicator saturation thus leads to an increase of  $\Delta[\text{Ca}^{2+}]_i$  per AP. The sharp increments of  $[\text{Ca}^{2+}]_i$  for each spike were smoothed with an exponential rising onset function ( $\tau_{\text{on}} = 20 \text{ ms}$ ) similar to Equation 4. Finally, we transformed the  $[\text{Ca}^{2+}]_i$  trace to a  $\Delta F/F$  trace using equation 1, presuming the following reasonable parameter values for OGB-1:  $K_d = 250 \text{ nM}$ ,  $[\text{B}]_T = 50 \mu\text{M}$ ,  $\Delta F/F_{\text{max}} = 93\%$ . With these parameter settings, a single-AP evoked  $\Delta F/F$  transient from resting  $[\text{Ca}^{2+}]_i$  level was similar to the stereotype  $\Delta F/F$  transient described by Equation 4. Note that despite the increased  $\Delta[\text{Ca}^{2+}]_i$  at elevated  $[\text{Ca}^{2+}]_i$  levels the non-linear transformation between  $[\text{Ca}^{2+}]_i$  and  $\Delta F/F$  (Equation 1) has the effect that the  $\Delta F/F$ -increment per AP becomes small closer to saturation (see Figure 4B).

For both the linear (low firing rate) and non-linear (higher firing rates) case, we added Gaussian white noise with standard deviation  $SD_{\text{noise}}$  to the simulated  $\Delta F/F$  traces. We assumed a realistic range of signal-to-noise ratios (SNR) for AP-evoked calcium transients, where we defined SNR as:

$$\text{SNR} = \frac{A_{\text{peak}}}{SD_{\text{noise}}} \quad (10)$$

We verified the assumption of Gaussian noise by empirically determining the noise distribution from random-access calcium imaging data (OGB-1; 490 Hz scan rate) (Grewe et al., 2010) when no spike had occurred (as verified by simultaneous electrophysiology). Without exception, noise distributions could be well-approximated by fitting a Gaussian curve ( $r^2 = 0.96 \pm 0.02$ ),

suggesting that residual noise in two-photon calcium imaging indeed can be assumed normally distributed (**Figures 9C,D**) and contains little, if any, auto-correlation at lags  $>0.1$  s (**Figure 9E**). Gaussian noise is a reasonable assumption because the number of detected photons is likely to be much greater than 100 under two-photon imaging conditions (Ranganathan and Koester, 2010). We note that for extremely low light conditions this assumption may not be valid. As the last step in our generation of simulated  $\Delta F/F$  traces, we subsampled the resulting noisy  $\Delta F/F$  trace from the original temporal resolution of 2 kHz to a given target frame rate,  $f$ , by selecting the center data point for each time interval  $\Delta t$ , where  $\Delta t = 1/f$ .

In summary, our analysis indicates that the presented simulation framework provides a valid model for AP-evoked calcium signals measured *in vivo* using two-photon microscopy. While experimental data may be characterized by additional noise sources not captured in our model (for example slow drifts or motion artifacts), these are generally easy to identify and remove prior to further data analysis. Whereas the linear description is appropriate for many cases and has been widely adopted (Yaksi and Friedrich, 2006; Vogelstein et al., 2010; Mishchenko et al., 2011), we have here also generalized our approach to the non-linear regime by considering indicator saturation. Extension to include further non-linearities—such as for example saturation of endogenous buffers, cooperative indicator  $\text{Ca}^{2+}$ -binding, e.g., for GECIs (Pologruto et al., 2004; Horikawa et al., 2010; Chen et al., 2013b), or diffusional equilibration—will be straight forward. Likewise, other non-linear relationships between  $[\text{Ca}^{2+}]_i$  and fluorescence readouts different from  $\Delta F/F$ , for example using ratiometric measurements, could also be considered.

## RECONSTRUCTION OF SPIKE TRAINS FROM CALCIUM INDICATOR SIGNALS

Action potentials were recovered from simulated  $\Delta F/F$  traces using the peeling algorithm that we have introduced previously (Grewe et al., 2010). Briefly, AP-evoked fluorescence signal events were detected using Schmitt-trigger thresholding (high threshold:  $+1.75$  SD, low threshold:  $-1$  SD, minimal duration: 0.3 s) with additional integral check (at least 50% of theoretical noise-free integral). In the original peeling algorithm we assumed a linear relationship between  $[\text{Ca}^{2+}]_i$  and  $\Delta F/F$ , which we also applied here for the low firing regime. Specifically, a stereotypical single-AP evoked  $\Delta F/F$  transient waveform (with the same parameters as used for the simulation of  $[\text{Ca}^{2+}]_i$  transients, unless noted otherwise) was iteratively subtracted (“peeled off”) as long as the integral of the residual trace remained positive and threshold-passing occurred.

An advantage of the model-based nature of the peeling algorithm is that a non-linearity like indicator saturation can be easily incorporated. Here, we extended the peeling algorithm to take saturation into account, in order to enable spike reconstruction from saturating  $\Delta F/F$  traces at high AP firing rates (up to 30 Hz). To this end, the single-AP evoked  $\Delta F/F$  transient was recalculated for each AP taking the respective pre-AP  $[\text{Ca}^{2+}]_i$  level into account (again presuming parameter values for OGB-1; see above). More specific, the  $[\text{Ca}^{2+}]_i$ -level dependent  $\Delta F/F$  transient was calculated by taking the difference between the  $\Delta F/F$

relaxation traces from post-AP and pre-AP levels (both computed by transforming the respective  $[\text{Ca}^{2+}]_i$  decays, obtained by solving the differential Equation 8). For comparison of error rates we applied either the simple linear or the saturating peeling algorithm to  $[\text{Ca}^{2+}]_i$  traces generated with a saturating indicator.

For both the linear and saturating peeling approach, the temporal precision of detected spikes was further improved by optimization of spike times ( $\pm 1$  s around the spike time determined with the peeling algorithm;  $\pm 0.1$  s for high AP rates). Optimization was performed by minimizing the squared sum of the residual trace using a pattern search algorithm (implemented in the Matlab Optimization toolbox).

To examine spike detection performance independent of the particular Schmitt-trigger thresholds, we performed “precision-recall” (PR) analysis (see **Table 1**) by selecting combinations of Schmitt-trigger thresholds over wide ranges (high threshold:  $-2$  to  $+5$  SD; low threshold:  $-5$  to  $+2$  SD; minimal duration: 0–1 s) (**Figures 2D,E**). Within the framework of PR analysis (Davis and Goadrich, 2006), we defined the break-even point as the data point closest to the unity line. Error rate  $\alpha_{\text{AP}}$  was defined as  $\max(\text{FDR}, 1 - \text{TPR})$  at this point (range 0–1). Intuitively,  $\alpha_{\text{AP}}$  describes the distance of the break-even point from the upper-right corner of the PR-curve, which represents optimal performance (Davis and Goadrich, 2006).

## COMPARISON OF ORIGINAL AND RECONSTRUCTED SPIKE TRAINS

Spike time comparison was performed by successively matching spikes in the original and reconstructed spike train based on ascending spike time difference (up to a maximal difference of 0.5 s, see **Figure 1D**). Remaining spikes in the original spike train reduce the true positive rate (calculated as fraction of total spikes in the original spike train) while spikes remaining in the reconstructed train contribute to the false discovery rate (calculated as a fraction of total spikes in the reconstructed spike

**Table 1 | Overview of spike metrics to quantify spike reconstruction accuracy.**

	Simulated spike train		
	Spike occurred	No spike occurred	Total
Reconstructed spike train			
Spike detected	True positive, TP	False discovery, FD	TP + FD
No spike detected	Missed spike, MS	Correct reject, CR	
Total	TP + MS		

Simulated spikes may be either detected (true positive, TP) or missed (MS) by the reconstruction algorithm. We call the ratio of true discoveries to the total number of simulated spikes the True Positive Rate,  $\text{TPR} = \text{TP}/(\text{TP} + \text{MS})$  or “recall.” On the other hand, spikes detected by the reconstruction algorithm may be matched by a simulated spike (true positive, TP) or represent false detections (false discovery, FD). We call the ratio of false discoveries to the total number of detected spikes the False Discovery Rate,  $\text{FDR} = \text{FD}/(\text{TP} + \text{FD})$ . In information retrieval theory (Davis and Goadrich, 2006),  $(1 - \text{FDR})$  is also known as “precision.” Both TPR and FDR are defined as 0 for the special case of no simulated or reconstructed spikes, respectively.



train). We quantify reconstruction performance by the following parameters (Table 1):

1. True positive rate (TPR): fraction of correctly detected spikes (out of total spikes in original spike train);  $TPR \in [0, 1]$ ,
2. False discovery rate (FDR): fraction of false discoveries (out of total spikes in reconstructed spike train);  $FDR \in [0, 1]$ ,
3. Temporal precision: mean and standard deviation of spike time differences between original and reconstructed spike trains,  $\text{mean}\Delta_t$  and  $\sigma_{\Delta_t}$ , respectively (only for correct detections).

## LARGE-SCALE NETWORK SIMULATION AND DETAILED NEURON MODEL

We simulated a network of 25,000 leaky integrate-and-fire neurons with conductance-based synapses (Zenke et al., 2013). 80% of the neurons were modeled as excitatory and 20% as inhibitory. Connectivity was chosen randomly with a density of 10%. In addition, each neuron received common excitatory input from a pool of 2000 independent Poisson processes that were connected randomly to all neurons with 10% probability. The rate of the external input was modeled as a pink noise stochastic process with a mean firing rate of 2 Hz per process and exhibiting fluctuations on all time scales ( $1/f$  power spectrum) to mimic complex temporal dynamics of common-input in cortical networks. The network was tuned to the balanced state with asynchronous and irregular firing activity with a mean spiking activity of  $\sim 0.2$  Hz.

Specifically, the membrane voltage  $U_i$  of a single cell  $i$  evolved according to:

$$\tau^m \frac{dU_i}{dt} = (U^{\text{rest}} - U_i) + g_i^{\text{exc}}(t) (U^{\text{exc}} - U_i) + g_i^{\text{inh}}(t) (U^{\text{inh}} - U_i) \quad (11)$$

with membrane time constants  $\tau^m = 20$  ms for excitatory neurons and  $\tau^m = 10$  ms for inhibitory neurons, resting potential  $U^{\text{rest}} = -70$  mV, reversal potentials  $U^{\text{exc}} = 0$  mV and  $U^{\text{inh}} = -80$  mV and conductances  $g_i^{\text{exc}}(t)$  and  $g_i^{\text{inh}}(t)$  specified below. A spike was triggered when  $U_i$  crossed the spiking threshold  $\vartheta_i$ . After each spike,  $U_i$  was reset to the resting value  $U^{\text{rest}}$  and the threshold  $\vartheta_i$  set to  $\vartheta^{\text{spike}} = 50$  mV to implement a refractoriness mechanism. Following a reset, the threshold exponentially decayed to its resting value  $\vartheta^{\text{rest}} = -50$  mV according to

$$\tau^{\text{thr}} \frac{d\vartheta_i}{dt} = (\vartheta^{\text{rest}} - \vartheta_i) \quad (12)$$

with time constant  $\tau^{\text{thr}} = 5$  ms. The spike train  $S_j(t)$  emitted by neuron  $j$  is given as  $S_j(t) = \sum_k \delta(t - t_j^k)$ , where the sum runs over all  $k$  corresponding spike times  $t_j^k$ . Inhibitory synaptic conductances of the downstream neurons were affected by presynaptic spikes as:

$$\tau^{\text{GABA}} \frac{dg_i^{\text{inh}}}{dt} = -g_i^{\text{inh}} + \sum_{j \in \text{inh}} w_{ij} S_j(t) \quad (13)$$

with  $\tau^{\text{GABA}} = 10$  ms. Excitatory synapses were modeled containing a fast AMPA component with exponential decay ( $\tau^{\text{AMPA}} =$

5 ms) and a slow NMDA component ( $\tau^{\text{NMDA}} = 100$  ms):

$$\tau^{\text{AMPA}} \frac{dg_i^{\text{AMPA}}}{dt} = -g_i^{\text{AMPA}} + \sum_{j \in \text{exc}} w_{ij} S_j(t) \quad (14)$$

$$\tau^{\text{NMDA}} \frac{dg_i^{\text{NMDA}}}{dt} = -g_i^{\text{NMDA}} + g_i^{\text{AMPA}} \quad (15)$$

The complete excitatory postsynaptic potential (EPSP) was obtained by a weighted sum of the AMPA and NMDA conductances:

$$g_i^{\text{exc}}(t) = 0.5g_i^{\text{AMPA}}(t) + 0.5g_i^{\text{NMDA}}(t) \quad (16)$$

The weight values  $w_{ij}$  of the synapse connecting neuron  $j$  with  $i$  ( $w_{ij} = 0$  if the connection does not exist) are given as follows:  $w(E \rightarrow E) = w(E \rightarrow I) = 0.2$  and  $w(I \rightarrow E) = w(I \rightarrow I) = 0.9$ . The external Poisson inputs were connected with a constant weight  $w(\text{ext} \rightarrow E, I) = 0.22$ . For computational efficiency, the voltage dependence of NMDA channels was omitted. All differential equations were integrated numerically using a forward Euler scheme with 0.1 ms time step using custom-written C/C++ code. Spike trains were generated for a total duration of  $T = 10,000$  s.

## CONNECTIVITY RECONSTRUCTION BASED ON COUPLED POINT PROCESS MODELS

We selected subsets of  $N = 50$  excitatory neurons from the population that had an average firing rate of 0.6 Hz or higher and reconstructed the connectivity between neurons of this subpopulation based on their spike trains of length  $T = 10,000$  s. To extract the coupling, we fitted coupled GLMs to the spike trains. Full details on the methodology can be found in (Gerhard et al., 2011). Briefly, spike trains are discretized into a sequence of binary values which represent spiking activity within time windows of length 1 ms. The instantaneous firing probability for each time bin is modeled as a non-linear transformation of the sum of covariates. These include effects from past spiking of the neuron itself as well as spikes from other neurons. All coupling filters are parameterized using a set of spline basis functions and parameters are estimated using standard maximum-likelihood techniques. Note that the strength of the stochastic common-input to each neuron is unobserved and therefore not explicitly modeled. The coefficients corresponding to the cross-coupling filters are used to define the effective coupling structure: The integral of each interaction filter represents its strength (Gerhard et al., 2013). A binary decision about the presence of a directed link can be enforced by thresholding the matrix of coupling strengths. The pair of  $\text{TPR}_{\text{links}}$  (fraction of correctly identified connections) and  $\text{FDR}_{\text{links}}$  (false discovery rate) defines the error rate for the link reconstruction as the smallest  $\alpha_{\text{links}}$  that guarantees  $\text{FDR}_{\text{links}} \leq \alpha_{\text{links}}$  and  $\text{TPR}_{\text{links}} \geq 1 - \alpha_{\text{links}}$ . Results generally show the averaged performance derived from the analysis of several random subpopulations of the full network. To derive the expected error rate in the link reconstruction under the assumption that the effect of the absolute detection power ( $\alpha_{\text{AP}}$ ) and spike time jitter ( $\sigma_{\Delta t}$ ) act independently (Figure 6F), we use the intuition that detection powers  $\sim 1 - \alpha$  would combine multiplicatively, so that,

approximately:

$$\alpha_{\text{links, independent}} \approx 1 - \frac{(1 - \alpha_{\text{links, due to } \alpha_{\text{AP}}})(1 - \alpha_{\text{links, due to } \sigma_{\Delta t}})}{1 - \alpha_{\text{links}}^*} \quad (17)$$

where  $\alpha_{\text{links}}^*$  is the best achievable error rate (in case of perfect spike reconstruction).

### CROSS-CORRELATION ANALYSIS

For comparison, we also implemented a connectivity extraction algorithm based on spike count correlations. We binned the spike trains into bins of size  $\Delta t_{\text{cc}}$  and calculated the pairwise Pearson's cross-correlation coefficient of the resulting time series for each pair of neurons in the selected subpopulation. The negative logarithm of the significance value, i.e., the surprise, served as coupling strength. Note that this yielded symmetric (i.e., bidirectional) couplings. We swept through a wide range of values for  $\Delta t_{\text{cc}}$  (0.5–500 ms) and chose the one with best performance, resulting in  $\Delta t_{\text{cc}} = 5$  ms.

### SURROGATE MODEL OF SPIKE TRAIN RECONSTRUCTION

We perturbed the spike trains using surrogate transformations to simulate the effect of the errors introduced by imperfect spike reconstruction from noisy calcium imaging data. Specifically, we used the two key parameters that were used to describe the performance of the single-neuron spike reconstruction (error rate  $\alpha_{\text{AP}}$  and spike jitter  $\sigma_{\Delta t}$ ). For any error rate  $\alpha_{\text{AP}} > 0$ , spikes were randomly removed from the simulated spike trains to match the desired  $\text{TPR}_{\text{AP}}$ . Simultaneously, spikes were added at random times up to the prescribed level of  $\text{FDR}_{\text{AP}}$ . The temporal imprecision  $\sigma_{\Delta t}$  was introduced by an additional jitter to all spike times given by a Gaussian distribution around zero with standard deviation  $\sigma_{\Delta t}$ . We repeated the connectivity estimation based on the perturbed spike trains and measured the performance using the error rate  $\alpha_{\text{links}}$  whose value should be compared to the reference value achievable in the case of unperturbed original spike trains (assuming perfect spike time reconstruction).

### IDENTIFICATION OF GRAPH TOPOLOGY

#### Scale-free networks

We generated scale-free networks of size 1000 neurons by constructing unweighted, undirected graphs whose degree distributions follow a power law  $p(x) \sim x^{-\mu}$  above a minimal degree  $k = 20$  with exponent  $\mu = 3$ , using the standard configuration model (Molloy and Reed, 1995).  $k$  was chosen as to produce an average link density of 4%, unless otherwise noted. We simulated the joint effect of calcium dynamics, spike train reconstruction and connectivity extraction by assuming that links are reconstructed with an error rate  $\alpha_{\text{links}}$ . This surrogate keeps the overall link density approximately constant. We then obtained the degree distribution of the reconstructed network and fitted a power law on its tail where the minimal degree and exponent were obtained using maximum-likelihood methods (Clauset et al., 2009). We constrained the exponent  $\mu$  to be between 1 and 9 which covers all empirically observed scale-free networks. A goodness-of-fit test was applied to each fit using a Monte Carlo version of the

Kolmogorov–Smirnov test (Clauset et al., 2009). We repeated the process of generation, imperfect reconstruction and re-fitting 1000 times and reported the median of the estimated power-law coefficient together with its standard deviation. We concluded that an estimated degree distribution was inconsistent with a power-law shape whenever the median  $p$ -value of the fit was below 0.05, i.e., a  $p$ -value  $< 0.05$  occurred in more than half of the cases. Histograms of degree distributions were obtained with logarithmically spaced bins and by pooling distributions across all simulations.

#### Hub neurons

We generated scale-free networks of size 1000 neurons and power-law exponent  $\mu = 3$  as described above. We classified hub neurons in these networks as the 100 neurons with the highest degrees. We then simulated an imperfect network reconstruction as before and estimated how well neurons can be classified to be hub neurons as follows: The hit rate specifies the fraction of original hub neurons that belong to the 100 neurons with highest degree in the reconstructed network. A random assignment would lead to a hit rate of 10% (chance level). All estimates are based on 1000 simulations of independent networks.

### ACKNOWLEDGMENTS

This work was supported by a grant from the Forschungskredit of the University of Zurich to Henry Lütcke; two grants from the Swiss National Science Foundation (SNSF) to Felipe Gerhard and Wulfram Gerstner (200020-117975 and 200020-132871); an SNSF grant to Fritjof Helmchen (3100A0-114624); the EU-FP7 program (BRAIN-I-NETS project 243914 and BrainScaleS project 269921 to Felipe Gerhard, Friedemann Zenke, Fritjof Helmchen, and Wulfram Gerstner) and a grant from the Swiss SystemsX.ch initiative (project 2008/2011-Neurochoice) to Fritjof Helmchen. We thank B. Grewe for collecting the data shown in Figures 9A,C.

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fncir.2013.00201/abstract>

### REFERENCES

- Badel, L., Lefort, S., Brette, R., Petersen, C. C., Gerstner, W., and Richardson, M. J. (2008). Dynamic I-V curves are reliable predictors of naturalistic pyramidal-neuron voltage traces. *J. Neurophysiol.* 99, 656–666. doi: 10.1152/jn.01107.2007
- Barabasi, A. L., and Albert, R. (1999). Emergence of scaling in random networks. *Science* 286, 509–512. doi: 10.1126/science.286.5439.509
- Bock, D. D., Lee, W. C., Kerlin, A. M., Andermann, M. L., Hood, G., Wetzell, A. W., et al. (2011). Network anatomy and *in vivo* physiology of visual cortical neurons. *Nature* 471, 177–182. doi: 10.1038/nature09802
- Bonin, V., Histed, M. H., Yurgenson, S., and Reid, R. C. (2011). Local diversity and fine-scale organization of receptive fields in mouse visual cortex. *J. Neurosci.* 31, 18506–18521. doi: 10.1523/JNEUROSCI.2974-11.2011
- Briggman, K. L., Helmstaedter, M., and Denk, W. (2011). Wiring specificity in the direction-selectivity circuit of the retina. *Nature* 471, 183–188. doi: 10.1038/nature09818
- Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.* 8, 183–208. doi: 10.1023/A:1008925309027

- Buzsaki, G. (2004). Large-scale recording of neuronal ensembles. *Nat. Neurosci.* 7, 446–451. doi: 10.1038/nn1233
- Chen, J. L., Carta, S., Soldado-Magraner, J., Schneider, B. L., and Helmchen, F. (2013a). Behavior-dependent recruitment of long-range projection neurons in somatosensory cortex. *Nature* 499, 336–340. doi: 10.1038/nature12236
- Chen, T. W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., et al. (2013b). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499, 295–300. doi: 10.1038/nature12354
- Chen, X., Leischner, U., Varga, Z., Jia, H., Deca, D., Rochefort, N. L., et al. (2012). LOTOS-based two-photon calcium imaging of dendritic spines *in vivo*. *Nat. Protoc.* 7, 1818–1829. doi: 10.1038/nprot.2012.106
- Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Rev.* 51, 661–703. doi: 10.1137/070710111
- Davis, J., and Goadrich, M. (2006). “The relationship between Precision-Recall and ROC curves,” in *Proceedings of the 23rd International Conference on Machine Learning*. (Pittsburgh, PA: ACM).
- Decharms, R. C., Blake, D. T., and Merzenich, M. M. (1998). Optimizing sound features for cortical neurons. *Science* 280, 1439–1443. doi: 10.1126/science.280.5368.1439
- De Kock, C. P., and Sakmann, B. (2008). High frequency action potential bursts ( $>100$  Hz) in L2/3 and L5B thick tufted neurons in anesthetized and awake rat primary somatosensory cortex. *J. Physiol.* 586, 3353–3364. doi: 10.1113/jphysiol.2008.155580
- Dombeck, D. A., Graziano, M. S., and Tank, D. W. (2009). Functional clustering of neurons in motor cortex determined by cellular resolution imaging in awake behaving mice. *J. Neurosci.* 29, 13751–13760. doi: 10.1523/JNEUROSCI.2985-09.2009
- Einevoll, G. T., Franke, F., Hagen, E., Pouzat, C., and Harris, K. D. (2011). Towards reliable spike-train recordings from thousands of neurons with multielectrodes. *Curr. Opin. Neurobiol.* 22, 11–17. doi: 10.1016/j.conb.2011.10.001
- Engelbrecht, C. J., Gobel, W., and Helmchen, F. (2009). Enhanced fluorescence signal in nonlinear microscopy through supplementary fiber-optic light collection. *Opt. Express* 17, 6421–6435. doi: 10.1364/OE.17.006421
- Feldt, S., Bonifazi, P., and Cossart, R. (2011). Dissecting functional connectivity of neuronal microcircuits: experimental and theoretical insights. *Trends Neurosci.* 34, 225–236. doi: 10.1016/j.tins.2011.02.007
- Fernández-Alfonso, T., Nadella, K. M., Iacuruso, M. F., Pichler, B., Roš, H., Kirkby, P. A., et al. (2013). Monitoring synaptic and neuronal activity in 3D with synthetic and genetic indicators using a compact acousto-optic lens two-photon microscope. *J. Neurosci. Methods* 222C, 69–81. doi: 10.1016/j.jneumeth.2013.10.021
- Fino, E., and Yuste, R. (2011). Dense inhibitory connectivity in neocortex. *Neuron* 69, 1188–1203. doi: 10.1016/j.neuron.2011.02.025
- Garaschuk, O., Milos, R. I., Grienberger, C., Marandi, N., Adelsberger, H., and Konnerth, A. (2006). Optical monitoring of brain function *in vivo*: from neurons to networks. *Pflügers Arch.* 453, 385–396. doi: 10.1007/s00424-006-0150-x
- Gerhard, F., Kispersky, T., Gutierrez, G. J., Marder, E., Kramer, M., and Eden, U. (2013). Successful reconstruction of a physiological circuit with known connectivity from spiking activity alone. *PLoS Comput. Biol.* 9:e1003138. doi: 10.1371/journal.pcbi.1003138
- Gerhard, F., Pipa, G., Lima, B., Neuenschwander, S., and Gerstner, W. (2011). Extraction of network topology from multi-electrode recordings: is there a small-world effect? *Front. Comput. Neurosci.* 5:4. doi: 10.3389/fncom.2011.00004
- Göbel, W., and Helmchen, F. (2007). *In vivo* calcium imaging of neural network function. *Physiology (Bethesda)* 22, 358–365. doi: 10.1152/physiol.00032.2007
- Greenberg, D. S., Houweling, A. R., and Kerr, J. N. (2008). Population imaging of ongoing neuronal activity in the visual cortex of awake rats. *Nat. Neurosci.* 11, 749–751. doi: 10.1038/nn.2140
- Grewe, B. F., and Helmchen, F. (2009). Optical probing of neuronal ensemble activity. *Curr. Opin. Neurobiol.* 19, 520–529. doi: 10.1016/j.conb.2009.09.003
- Grewe, B. F., Langer, D., Kasper, H., Kampa, B. M., and Helmchen, F. (2010). High-speed *in vivo* calcium imaging reveals neuronal network activity with near-millisecond precision. *Nat. Methods* 7, 399–405. doi: 10.1038/nmeth.1453
- Grienberger, C., and Konnerth, A. (2012). Imaging calcium in neurons. *Neuron* 73, 862–885. doi: 10.1016/j.neuron.2012.02.011
- Han, J. D., Dupuy, D., Bertin, N., Cusick, M. E., and Vidal, M. (2005). Effect of sampling on topology predictions of protein-protein interaction networks. *Nat. Biotechnol.* 23, 839–844. doi: 10.1038/nbt1116
- Helmchen, F. (2012). “Calcium imaging,” in *Handbook of Neural Activity Measurement*, eds R. Brette and A. Destexhe (Cambridge: Cambridge University Press).
- Helmchen, F., Imoto, K., and Sakmann, B. (1996).  $\text{Ca}^{2+}$  buffering and action potential-evoked  $\text{Ca}^{2+}$  signaling in dendrites of pyramidal neurons. *Biophys. J.* 70, 1069–1081. doi: 10.1016/S0006-3495(96)79653-4
- Helmchen, F., and Tank, D. W. (2011). “A single compartment model of calcium dynamics in dendrites and nerve terminals,” in *Imaging in Neuroscience: A Laboratory Manual*, eds F. Helmchen and A. Konnerth (New York, NY: Cold Spring Harbor Laboratory Press).
- Hendel, T., Mank, M., Schnell, B., Griesbeck, O., Borst, A., and Reiff, D. F. (2008). Fluorescence changes of genetic calcium indicators and OGB-1 correlated with neural activity and calcium *in vivo* and *in vitro*. *J. Neurosci.* 28, 7399–7411. doi: 10.1523/JNEUROSCI.1038-08.2008
- Hofer, S. B., Ko, H., Pichler, B., Vogelstein, J., Ros, H., Zeng, H., et al. (2011). Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nat. Neurosci.* 14, 1045–1052. doi: 10.1038/nn.2876
- Horikawa, K., Yamada, Y., Matsuda, T., Kobayashi, K., Hashimoto, M., Matsuura, T., et al. (2010). Spontaneous network activity visualized by ultrasensitive  $\text{Ca}^{2+}$  indicators, yellow Cameleon-Nano. *Nat. Methods* 7, 729–732. doi: 10.1038/nmeth.1488
- Kampa, B. M., Göbel, W., and Helmchen, F. (2011). Measuring neuronal population activity using 3D laser scanning. *Cold Spring Harb. Protoc.* 2011, 1340–1349. doi: 10.1101/pdb.prot066597
- Katona, G., Szalay, G., Maak, P., Kaszas, A., Veress, M., Hillier, D., et al. (2012). Fast two-photon *in vivo* imaging with three-dimensional random-access scanning in large tissue volumes. *Nat. Methods* 9, 201–208. doi: 10.1038/nmeth.1851
- Kerlin, A. M., Andermann, M. L., Berezovskii, V. K., and Reid, R. C. (2010). Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* 67, 858–871. doi: 10.1016/j.neuron.2010.08.002
- Kerr, J. N., De Kock, C. P., Greenberg, D. S., Bruno, R. M., Sakmann, B., and Helmchen, F. (2007). Spatial organization of neuronal population responses in layer 2/3 of rat barrel cortex. *J. Neurosci.* 27, 13316–13328. doi: 10.1523/JNEUROSCI.2210-07.2007
- Kerr, J. N., Greenberg, D., and Helmchen, F. (2005). Imaging input and output of neocortical networks *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14063–14068. doi: 10.1073/pnas.0506029102
- Knopfel, T. (2012). Genetically encoded optical indicators for the analysis of neuronal circuits. *Nat. Rev. Neurosci.* 13, 687–700. doi: 10.1038/nrn3293
- Ko, H., Hofer, S. B., Pichler, B., Buchanan, K. A., Sjöström, P. J., and Mrisic-Flogel, T. D. (2011). Functional specificity of local synaptic connections in neocortical networks. *Nature* 473, 87–91. doi: 10.1038/nature09880
- Langer, D., and Helmchen, F. (2012). *Post hoc* immunostaining of GABAergic neuronal subtypes following *in vivo* two-photon calcium imaging in mouse neocortex. *Pflügers Arch.* 463, 339–354. doi: 10.1007/s00424-011-1048-9
- Lillis, K. P., Eng, A., White, J. A., and Mertz, J. (2008). Two-photon imaging of spatially extended neuronal network dynamics with high temporal resolution. *J. Neurosci. Methods* 172, 178–184. doi: 10.1016/j.jneumeth.2008.04.024
- Looger, L. L., and Griesbeck, O. (2012). Genetically encoded neural activity indicators. *Curr. Opin. Neurobiol.* 22, 18–23. doi: 10.1016/j.conb.2011.10.024
- Lütcke, H., and Helmchen, F. (2011). Two-photon imaging and analysis of neural network dynamics. *Rep. Prog. Phys.* 74. doi: 10.1088/0034-4885/74/8/086602
- Lütcke, H., Margolis, D. J., and Helmchen, F. (2013). Steady or changing? Long-term monitoring of neuronal population activity. *Trends Neurosci.* 36, 375–384. doi: 10.1016/j.tins.2013.03.008
- Lütcke, H., Murayama, M., Hahn, T., Margolis, D. J., Astori, S., Zum Alten Borgloh, S. M., et al. (2010). Optical recording of neuronal activity with a genetically encoded calcium indicator in anesthetized and freely moving mice. *Front. Neural Circuits* 4:9. doi: 10.3389/fncir.2010.00009
- Margolis, D. J., Lütcke, H., Schulz, K., Haiss, F., Weber, B., Kügler, S., et al. (2012). Reorganization of cortical population activity imaged throughout long-term sensory deprivation. *Nat. Neurosci.* 15, 1539–1546. doi: 10.1038/nn.3240
- Mensi, S., Naud, R., Pozzorini, C., Avermann, M., Petersen, C. C., and Gerstner, W. (2012). Parameter extraction and classification of three cortical neuron types reveals two distinct adaptation mechanisms. *J. Neurophysiol.* 107, 1756–1775. doi: 10.1152/jn.00408.2011
- Mishchenko, Y., Vogelstein, J. T., and Paninski, L. (2011). A bayesian approach for inferring neuronal connectivity from calcium fluorescent imaging data. *Ann. Appl. Stat.* 5, 1229–1261. doi: 10.1214/09-AOAS303

- Molloy, M., and Reed, B. (1995). A critical-point for random graphs with a given degree sequence. *Random Struct. Algor.* 6, 161–179. doi: 10.1002/rsa.3240060204
- Nikolenko, V., Poskanzer, K. E., and Yuste, R. (2007). Two-photon photostimulation and imaging of neural circuits. *Nat. Methods* 4, 943–950. doi: 10.1038/nmeth1105
- Nikolenko, V., Watson, B. O., Araya, R., Woodruff, A., Peterka, D. S., and Yuste, R. (2008). SLM microscopy: scanless two-photon imaging and photostimulation with spatial light modulators. *Front. Neural Circuits* 2:5. doi: 10.3389/neuro.04.005.2008
- Oheim, M., Beaupaire, E., Chaigneau, E., Mertz, J., and Charpak, S. (2001). Two-photon microscopy in brain tissue: parameters influencing the imaging depth. *J. Neurosci. Methods* 111, 29–37. doi: 10.1016/S0165-0270(01)00438-1
- Ohki, K., Chung, S., Ch'ng, Y., Kara, P., and Reid, R. (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* 433, 597–603. doi: 10.1038/nature03274
- Ohki, K., Chung, S., Kara, P., Hubener, M., Bonhoeffer, T., and Reid, R. C. (2006). Highly ordered arrangement of single neurons in orientation pinwheels. *Nature* 442, 925–928. doi: 10.1038/nature05019
- Olshausen, B. A., and Field, D. J. (2005). How close are we to understanding v1? *Neural Comput.* 17, 1665–1699. doi: 10.1162/0899766054026639
- Onativia, J., Schultz, S. R., and Dragotti, P. L. (2013). A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging. *J. Neural Eng.* 10, 046017. doi: 10.1088/1741-2560/10/4/046017
- Osten, P., and Margrie, T. W. (2013). Mapping brain circuitry with a light microscope. *Nat. Methods* 10, 515–523. doi: 10.1038/nmeth.2477
- Otsu, Y., Bormuth, V., Wong, J., Mathieu, B., Dugue, G. P., Feltz, A., et al. (2008). Optical monitoring of neuronal activity at high frame rate with a digital random-access multiphoton (RAMP) microscope. *J. Neurosci. Methods* 173, 259–270. doi: 10.1016/j.jneumeth.2008.06.015
- Packer, A. M., and Yuste, R. (2011). Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *J. Neurosci.* 31, 13260–13271. doi: 10.1523/JNEUROSCI.3131-11.2011
- Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., et al. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454, 995–999. doi: 10.1038/nature07140
- Pologruto, T. A., Yasuda, R., and Svoboda, K. (2004). Monitoring neural activity and [Ca<sup>2+</sup>] with genetically encoded Ca<sup>2+</sup> indicators. *J. Neurosci.* 24, 9572–9579. doi: 10.1523/JNEUROSCI.2854-04.2004
- Prakash, R., Yizhar, O., Grewe, B., Ramakrishnan, C., Wang, N., Goshen, I., et al. (2012). Two-photon optogenetic toolbox for fast inhibition, excitation and bistable modulation. *Nat. Methods* 9, 1171–1179. doi: 10.1038/nmeth.2215
- Ranganathan, G. N., and Koester, H. J. (2010). Optical recording of neuronal spiking activity from unbiased populations of neurons with high spike detection efficiency and high temporal precision. *J. Neurophysiol.* 104, 1812–1824. doi: 10.1152/jn.00197.2010
- Reddy, G. D., and Saggau, P. (2005). Fast three-dimensional laser scanning scheme using acousto-optic deflectors. *J. Biomed. Opt.* 10, 064038. doi: 10.1117/1.2141504
- Rocheft, N. L., Garaschuk, O., Milos, R. I., Narushima, M., Marandi, N., Pichler, B., et al. (2009). Sparsification of neuronal activity in the visual cortex at eye-opening. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15049–15054. doi: 10.1073/pnas.0907660106
- Rothschild, G., Nelken, I., and Mizrahi, A. (2010). Functional organization and population dynamics in the mouse primary auditory cortex. *Nat. Neurosci.* 13, 353–360. doi: 10.1038/nn.2484
- Sasaki, T., Takahashi, N., Matsuki, N., and Ikegaya, Y. (2008). Fast and accurate detection of action potentials from somatic calcium fluctuations. *J. Neurophysiol.* 100, 1668–1676. doi: 10.1152/jn.00084.2008
- Sjulson, L., and Miesenbock, G. (2007). Optical recording of action potentials and other discrete physiological events: a perspective from signal detection theory. *Physiology* 22, 47–55. doi: 10.1152/physiol.00036.2006
- Smith, A. C., and Brown, E. N. (2003). Estimating a state-space model from point process observations. *Neural Comput.* 15, 965–991. doi: 10.1162/089976603765202622
- Song, S., Sjostrom, P. J., Reigl, M., Nelson, S., and Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* 3:e68. doi: 10.1371/journal.pbio.0030068
- Stetter, O., Battaglia, D., Soriano, J., and Geisel, T. (2012). Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals. *PLoS Comput. Biol.* 8:e1002653. doi: 10.1371/journal.pcbi.1002653
- Stumpf, M. P., Wiuf, C., and May, R. M. (2005). Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 4221–4224. doi: 10.1073/pnas.0501179102
- Turaga, S. C., Buesing, L., Packer, A. M., Dagleish, M., Petit, N., Haussler, M., et al. (in press). Inferring neural population dynamics from multiple partial recordings of the same neural circuit. *Adv. Neural Inf. Process. Syst.*
- Vidne, M., Ahmadian, Y., Shlens, J., Pillow, J. W., Kulkarni, J., Litke, A. M., et al. (2012). Modeling the impact of common noise inputs on the network activity of retinal ganglion cells. *J. Comput. Neurosci.* 33, 97–121. doi: 10.1007/s10827-011-0376-2
- Vogels, T. P., Rajan, K., and Abbott, L. F. (2005). Neural network dynamics. *Annu. Rev. Neurosci.* 28, 357–376. doi: 10.1146/annurev.neuro.28.061604.135637
- Vogelstein, J. T., Packer, A. M., Machado, T. A., Sipsey, T., Babadi, B., Yuste, R., et al. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. *J. Neurophysiol.* 104, 3691–3704. doi: 10.1152/jn.01073.2009
- Vogelstein, J. T., Watson, B. O., Packer, A. M., Yuste, R., Jedynek, B., and Paninski, L. (2009). Spike inference from calcium imaging using sequential Monte Carlo methods. *Biophys. J.* 97, 636–655. doi: 10.1016/j.bpj.2008.08.005
- Wilt, B. A., Fitzgerald, J. E., and Schnitzer, M. J. (2013). Photon shot noise limits on optical detection of neuronal spikes and estimation of spike timing. *Biophys. J.* 104, 51–62. doi: 10.1016/j.bpj.2012.07.058
- Wolfe, J., Houweling, A. R., and Brecht, M. (2010). Sparse and powerful cortical spikes. *Curr. Opin. Neurobiol.* 20, 306–312. doi: 10.1016/j.conb.2010.03.006
- Yaksi, E., and Friedrich, R. W. (2006). Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca<sup>2+</sup> imaging. *Nat. Methods* 3, 377–383. doi: 10.1038/nmeth874
- Zenke, F., Hennequin, G., and Gerstner, W. (2013). Synaptic plasticity in neural networks needs homeostasis with a fast rate detector. *PLoS Comput. Biol.* 9:e1003330. doi: 10.1371/journal.pcbi.1003330

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 July 2013; accepted: 04 December 2013; published online: 24 December 2013.

Citation: Lütcke H, Gerhard F, Zenke F, Gerstner W and Helmchen F (2013) Inference of neuronal network spike dynamics and topology from calcium imaging data. *Front. Neural Circuits* 7:201. doi: 10.3389/fncir.2013.00201

This article was submitted to the journal *Frontiers in Neural Circuits*.

Copyright © 2013 Lütcke, Gerhard, Zenke, Gerstner and Helmchen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.